## Contents

### Preface
- Audience xvii
- Documentation Accessibility xvii
- Related Documents xviii
- Conventions xviii

### 1. Introducing Oracle Exadata System Software

1.1 Overview of Oracle Exadata System Software 1-1
1.2 Key Features of Oracle Exadata System Software 1-2
  1.2.1 Reliability, Modularity, and Cost-Effectiveness 1-3
  1.2.2 Compatibility with Oracle Database 1-4
  1.2.3 Smart Flash Technology 1-4
    1.2.3.1 Exadata Smart Flash Cache 1-4
    1.2.3.2 Write-Back Flash Cache 1-5
    1.2.3.3 Exadata Smart Flash Log 1-5
  1.2.4 Persistent Memory Accelerator and RDMA 1-6
  1.2.5 Centralized Storage 1-7
  1.2.6 I/O Resource Management (IORM) 1-7
  1.2.7 Exadata Hybrid Columnar Compression 1-8
  1.2.8 In-Memory Columnar Format Support 1-11
  1.2.9 Offloading of Data Search and Retrieval Processing 1-12
  1.2.10 Offloading of Incremental Backup Processing 1-13
  1.2.11 Fault Isolation with Quarantine 1-13
    1.2.11.1 Quarantine Manager Support for Cell-to-Cell Offload Operations 1-15
  1.2.12 Protection Against Data Corruption 1-16
  1.2.13 Fast File Creation 1-17
  1.2.14 Storage Index 1-17
1.3 Oracle Exadata System Software Components 1-20
  1.3.1 About Oracle Exadata System Software 1-20
  1.3.2 About Oracle Automatic Storage Management 1-23
    1.3.2.1 Oracle ASM Disk Groups 1-23
    1.3.2.2 Oracle ASM Failure Group 1-24
1.3.2.3  Maximum Availability with Oracle ASM  1-25
1.3.3  About Grid RAID  1-28
1.3.4  About Storage Server Security  1-29
1.3.5  About iDB Protocol  1-29
1.3.6  About Oracle Exadata System Software Processes  1-29
1.3.7  About Cell Management  1-30
1.3.8  About Database Server Software  1-30
1.3.9  About Oracle Enterprise Manager for Oracle Exadata Database Machine  1-31

2  Configuring Oracle Exadata System Software

2.1  Understanding Oracle Exadata System Software Release Numbering  2-2
2.2  Understanding Oracle Exadata Storage Server Configuration  2-3
   2.2.1  Assign IP Addresses for the Storage Servers  2-3
   2.2.2  Configure the Storage Server for Your Location  2-4
   2.2.3  Configure the Storage Cell  2-4
   2.2.4  Verify Storage Cell Attributes  2-4
   2.2.5  Create the Storage Cell Disks  2-5
   2.2.6  Create the Grid Disks  2-6
   2.2.7  Create the PMEM Cache  2-8
   2.2.8  Create the Flash Disks and Flash Cache  2-8
   2.2.9  Configure Oracle Auto Service Request (ASR)  2-9
2.3  Network Configuration and IP Addresses Recommendations  2-9
2.4  Assigning IP Addresses for Oracle Exadata Database Machine  2-11
2.5  Configuring Oracle Exadata System Software for Your Location  2-12
   2.5.1  Configuring ILOM With Static IP for Oracle Exadata Storage Servers  2-12
   2.5.2  Preparing the Servers  2-13
2.6  Configuring Cells, Cell Disks, and Grid Disks with CellCLI  2-15
2.7  Creating Flash Cache and Flash Grid Disks  2-17
2.8  Setting Up Configuration Files for a Database Server Host  2-19
2.9  Understanding Automated Cell Maintenance  2-20

3  Administering Oracle ASM on Exadata

3.1  Overview of Oracle Exadata Database Machine Storage  3-1
3.2  Administering Oracle ASM Disk Groups Using Oracle Exadata Storage Servers  3-2
   3.2.1  Understanding Oracle ASM Disk Groups for Oracle Exadata Storage Servers  3-3
   3.2.1.1  About Fast Disk Scan Rates  3-4
   3.2.1.2  Setting the Oracle ASM Content Type  3-5
3.2.1.3 Creating Oracle ASM Disk Groups 3-6
3.2.2 Adding a Disk to an Oracle ASM Disk Group 3-9
3.2.3 Mounting or Dismounting an Oracle ASM Disk Group 3-9
3.2.4 Changing a Disk to Offline or Online 3-10
3.2.5 Dropping a Disk from an Oracle ASM Disk Group 3-10
3.2.6 Dropping an Oracle ASM Disk Group 3-11
3.2.7 Enabling the Oracle ASM appliance.mode Attribute 3-12
3.2.8 Checking Disk Group Balance 3-13
3.2.9 Setting the Oracle ASM Disk Repair Timer 3-13

3.3 Administering Oracle Exadata Storage Server Grid Disks with Oracle ASM 3-14
3.3.1 Naming Conventions for Oracle Exadata Storage Server Grid Disks 3-14
3.3.2 Changing an Oracle Exadata Storage Server Grid Disk That Belongs to an Oracle ASM Disk Group 3-15
   3.3.2.1 Changing an Oracle Exadata Storage Server Grid Disk Name 3-15
   3.3.2.2 Dropping an Oracle Exadata Storage Server Grid Disk 3-16
3.3.3 Resizing Grid Disks 3-17
   3.3.3.1 Determine the Amount of Available Space 3-19
   3.3.3.2 Shrink the Oracle ASM Disks in the Donor Disk Group 3-24
   3.3.3.3 Shrink the Grid Disks in the Donor Disk Group 3-26
   3.3.3.4 Increase the Size of the Grid Disks Using Available Space 3-27
   3.3.3.5 Increase the Size of the Oracle ASM Disks 3-29
3.3.4 Determining Which Oracle ASM Disk Group Contains an Oracle Exadata Storage Server Grid Disk 3-31
3.3.5 Determining Which Oracle Exadata Storage Server Grid Disks Belong to an Oracle ASM Disk Group 3-31
3.3.6 Handling Disk Replacement 3-32

4 Administering Oracle Database on Exadata

4.1 Administering SQL Processing Offload 4-1
   4.1.1 CELL_OFFLOAD_PROCESSING 4-1
   4.1.2 CELL_OFFLOAD_PLAN_DISPLAY 4-2
   4.1.3 CELL_OFFLOAD_DECRYPTION 4-3
4.2 Overriding the Default Caching Policy 4-3
4.3 Administering In-Memory Columnar Caching 4-4
4.4 Administering Exadata Hybrid Columnar Compression 4-5
   4.4.1 Determining If a Table Is Compressed 4-5
   4.4.2 Determining Which Rows are Compressed 4-6
   4.4.3 Changing Compression Level 4-7
   4.4.4 Importing and Exporting Exadata Hybrid Columnar Compression Tables 4-8
   4.4.5 Restoring an Exadata Hybrid Columnar Compression Table 4-8
4.5 Administering Oracle Database Features on Exadata 4-9
6.1.5 About Category Resource Management 6-12
6.2 About Consumer Groups and Resource Plans 6-13
   6.2.1 DSS_PLAN Resource Plan 6-13
   6.2.2 ETL_CRITCAL_PLAN Resource Plan 6-14
6.3 About CDB Plans and Pluggable Databases 6-14
6.4 Administering IORM 6-15
   6.4.1 Enabling IORM to Prioritize Latency 6-16
   6.4.2 Administering Database Resource Management 6-17
      6.4.2.1 Setting Up Consumer Groups and Categories 6-18
      6.4.2.2 Assigning Sessions to Consumer Groups 6-20
      6.4.2.3 Creating a CDB Plan 6-21
      6.4.2.4 Creating a Database Plan 6-22
      6.4.2.5 Enabling a Database Resource Plan 6-23
      6.4.2.6 Managing Fast File Creation 6-24
      6.4.2.7 Managing Data Import 6-24
      6.4.2.8 Managing Oracle Recovery Manager Backups and Copies 6-25
   6.4.3 Administering Interdatabase Resource Management 6-26
      6.4.3.1 How to Create an Interdatabase or Category Plan 6-26
      6.4.3.2 Required Settings for an Interdatabase or Category Plan 6-27
      6.4.3.3 Optional Settings for an Interdatabase or Category Plan 6-30
   6.4.4 Using Flash Cache Attributes in IORM Plans 6-31
   6.4.5 Managing Flash Cache Quotas for Databases and PDBs 6-32
   6.4.6 Using PMEM Cache Attributes in IORM Plans 6-34
   6.4.7 Managing PMEM Cache Quotas for Databases and PDBs 6-35
   6.4.8 Controlling Access to Flash Cache and Flash Log 6-37
   6.4.9 Controlling Access to PMEM Cache and PMEM Log 6-38
   6.4.10 Listing an I/O Resource Management Plan 6-39
   6.4.11 Verifying the Configuration of I/O Resource Management 6-39
   6.4.12 Resetting Default Values in an Interdatabase Plan 6-42

7 Monitoring Exadata

7.1 Introducing Exadata Monitoring Tools and Information Sources 7-1
    7.1.1 Automatic Workload Repository (AWR) 7-1
    7.1.2 Database Statistics and Wait Events 7-5
    7.1.3 Exadata Metrics 7-9
    7.1.4 Exadata Alerts 7-10
    7.1.5 ExaWatcher 7-11
7.2 Guidelines for Exadata Monitoring 7-11
7.3 Monitoring Oracle Exadata System Software Components 7-11
    7.3.1 Monitoring Exadata Smart Flash Cache 7-12
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3.1.1</td>
<td>Monitoring Exadata Smart Flash Cache Using AWR</td>
<td>7-13</td>
</tr>
<tr>
<td>7.3.1.2</td>
<td>Monitoring Exadata Smart Flash Cache Using Database Statistics and Wait Events</td>
<td>7-16</td>
</tr>
<tr>
<td>7.3.1.3</td>
<td>Monitoring Exadata Smart Flash Cache Using Exadata Metrics</td>
<td>7-18</td>
</tr>
<tr>
<td>7.3.1.4</td>
<td>What to Look For When Monitoring Exadata Smart Flash Cache</td>
<td>7-25</td>
</tr>
<tr>
<td>7.3.2.1</td>
<td>Monitoring PMEM Cache Using AWR</td>
<td>7-28</td>
</tr>
<tr>
<td>7.3.2.2</td>
<td>Monitoring PMEM Cache Using Database Statistics and Wait Events</td>
<td>7-30</td>
</tr>
<tr>
<td>7.3.2.3</td>
<td>Monitoring PMEM Cache Using Exadata Metrics</td>
<td>7-31</td>
</tr>
<tr>
<td>7.3.2.4</td>
<td>What to Look For When Monitoring PMEM Cache</td>
<td>7-32</td>
</tr>
<tr>
<td>7.3.3.1</td>
<td>Monitoring Exadata Smart Flash Log Using AWR</td>
<td>7-33</td>
</tr>
<tr>
<td>7.3.3.2</td>
<td>Monitoring Exadata Smart Flash Log Using Database Statistics and Wait Events</td>
<td>7-35</td>
</tr>
<tr>
<td>7.3.3.3</td>
<td>Monitoring Exadata Smart Flash Log Using Exadata Metrics</td>
<td>7-35</td>
</tr>
<tr>
<td>7.3.3.4</td>
<td>What to Look For When Monitoring Exadata Smart Flash Log</td>
<td>7-37</td>
</tr>
<tr>
<td>7.3.4.1</td>
<td>Monitoring PMEM Log Using Database Statistics and Wait Events</td>
<td>7-39</td>
</tr>
<tr>
<td>7.3.4.2</td>
<td>What to Look For When Monitoring PMEM Log</td>
<td>7-40</td>
</tr>
<tr>
<td>7.3.5.1</td>
<td>Monitoring Smart I/O Using AWR</td>
<td>7-41</td>
</tr>
<tr>
<td>7.3.5.2</td>
<td>Monitoring Smart I/O Using Database Statistics and Wait Events</td>
<td>7-42</td>
</tr>
<tr>
<td>7.3.5.3</td>
<td>Monitoring Smart I/O Using SQL Monitor</td>
<td>7-44</td>
</tr>
<tr>
<td>7.3.5.4</td>
<td>Monitoring Smart I/O Using SQL Explain Plan</td>
<td>7-46</td>
</tr>
<tr>
<td>7.3.5.5</td>
<td>Monitoring Smart I/O Using Exadata Metrics</td>
<td>7-48</td>
</tr>
<tr>
<td>7.3.5.6</td>
<td>What to Look For When Monitoring Smart I/O</td>
<td>7-50</td>
</tr>
<tr>
<td>7.3.6.1</td>
<td>Monitoring I/O Resource Management (IORM) Using AWR</td>
<td>7-54</td>
</tr>
<tr>
<td>7.3.6.2</td>
<td>Monitoring I/O Resource Management (IORM) Using Database Statistics</td>
<td>7-55</td>
</tr>
<tr>
<td>7.3.6.3</td>
<td>Monitoring I/O Resource Management (IORM) Using Exadata Metrics</td>
<td>7-56</td>
</tr>
<tr>
<td>7.3.6.4</td>
<td>What to Look For When Monitoring I/O Resource Management (IORM)</td>
<td>7-68</td>
</tr>
<tr>
<td>7.3.7.1</td>
<td>Monitoring Cell Disk I/O Using AWR</td>
<td>7-69</td>
</tr>
<tr>
<td>7.3.7.2</td>
<td>Monitoring Cell Disk I/O Using Database Statistics and Wait Events</td>
<td>7-74</td>
</tr>
<tr>
<td>7.3.7.3</td>
<td>Monitoring Cell Disk I/O Using Exadata Metrics</td>
<td>7-76</td>
</tr>
<tr>
<td>7.3.7.4</td>
<td>What to Look For When Monitoring Cell Disk I/O</td>
<td>7-78</td>
</tr>
<tr>
<td>7.3.8</td>
<td>Monitoring Grid Disks</td>
<td>7-80</td>
</tr>
<tr>
<td>7.3.9</td>
<td>Monitoring Host Interconnect Metrics</td>
<td>7-81</td>
</tr>
</tbody>
</table>
7.3.10 Monitoring RAM Cache
7.3.10.1 Monitoring RAM Cache Using AWR
7.3.10.2 Monitoring RAM Cache Using Database Statistics and Wait Events
7.4 Using Exadata Monitoring Objects
7.4.1 Displaying Metrics
7.4.1.1 Displaying Metric Definitions
7.4.1.2 Displaying Current Metrics
7.4.1.3 Displaying Metric History
7.4.2 Monitoring Alerts
7.4.2.1 Displaying Alert Definitions
7.4.2.2 Receiving Alert Notifications
7.4.2.3 Displaying Alert History
7.4.2.4 Modifying Alert History
7.4.3 Displaying Active Requests

8 Using the CellCLI Utility
8.1 Overview of the CellCLI Utility
8.1.1 Starting CellCLI
8.1.2 Understanding Command Syntax and Options for CellCLI
8.1.3 Reserved Words
8.1.4 CellCLI Command-Line Editing
8.1.5 CellCLI Input and Output Options
8.1.6 Comments in CellCLI Scripts
8.1.7 Line Continuation in CellCLI Commands
8.2 About CellCLI Administration Commands
8.3 About CellCLI Object Commands
8.4 About CellCLI Object Types
8.5 About Quoting Object Names
8.6 About CellCLI Object Attributes
8.6.1 Restrictions on Values of Common Attributes
8.6.2 Attribute Lists in LIST Command
8.6.3 Attribute Filters in LIST and ALTER GRIDDISK Commands
8.7 CellCLI Command Reference
8.7.1 ALTER
8.7.1.1 ALTER ALERTHISTORY
8.7.1.2 ALTER CELL
8.7.1.3 ALTER CELLDISK
8.7.1.4 ALTER FLASHCACHE
8.7.1.5 ALTER GRIDDISK
8.7.1.6 ALTER IBPORT
8.7.1.7 ALTER IORMPLAN 8-42
8.7.1.8 ALTER LUN 8-59
8.7.1.9 ALTER OFFLOADGROUP 8-60
8.7.1.10 ALTER PHYSICALDISK 8-60
8.7.1.11 ALTER PMEMCACHE 8-61
8.7.1.12 ALTER QUARANTINE 8-63
8.7.1.13 ALTER SOFTWAREUPDATE 8-63
8.7.1.14 ALTER THRESHOLD 8-66
8.7.1.15 ALTER USER 8-67

8.7.2 ASSIGN KEY 8-67
8.7.3 CALIBRATE 8-69

8.7.4 CREATE 8-71
8.7.4.1 CREATE CELL 8-72
8.7.4.2 CREATE CELLDISK 8-74
8.7.4.3 CREATE DIAGPACK 8-76
8.7.4.4 CREATE FLASHCACHE 8-80
8.7.4.5 CREATE FLASHLOG 8-81
8.7.4.6 CREATE GRIDDISK 8-83
8.7.4.7 CREATE KEY 8-86
8.7.4.8 CREATE PMEMCACHE 8-86
8.7.4.9 CREATE PMEMLOG 8-87
8.7.4.10 CREATE QUARANTINE 8-88
8.7.4.11 CREATE ROLE 8-89
8.7.4.12 CREATE THRESHOLD 8-90
8.7.4.13 CREATE USER 8-91

8.7.5 DESCRIBE 8-92
8.7.5.1 DESCRIBE ACTIVEREQUEST 8-93
8.7.5.2 DESCRIBE ALERTDEFINITION 8-95
8.7.5.3 DESCRIBE ALERTHISTORY 8-96
8.7.5.4 DESCRIBE CELL 8-97
8.7.5.5 DESCRIBE CELLDISK 8-103
8.7.5.6 DESCRIBE DATABASE 8-105
8.7.5.7 DESCRIBE DISKMAP 8-106
8.7.5.8 DESCRIBE FLASHCACHE 8-106
8.7.5.9 DESCRIBE FLASHCACHECONTENT 8-107
8.7.5.10 DESCRIBE FLASHLOG 8-109
8.7.5.11 DESCRIBE GRIDDISK 8-110
8.7.5.12 DESCRIBE IBPORT 8-112
8.7.5.13 DESCRIBE IORMPLAN 8-113
8.7.5.14 DESCRIBE KEY 8-114
8.7.5.15 DESCRIBE LUN 8-115
8.7.5.16  DESCRIBE METRICCURRENT 8-116
8.7.5.17  DESCRIBE METRICDEFINITION 8-117
8.7.5.18  DESCRIBE METRICHISTORY 8-120
8.7.5.19  DESCRIBE OFFLOADGROUP 8-122
8.7.5.20  DESCRIBE PHYSICALDISK 8-122
8.7.5.21  DESCRIBE PLUGGABLEDATABASE 8-125
8.7.5.22  DESCRIBE PMEMCACHE 8-126
8.7.5.23  DESCRIBE PMEMLOG 8-127
8.7.5.24  DESCRIBE QUARANTINE 8-128
8.7.5.25  DESCRIBE ROLE 8-129
8.7.5.26  DESCRIBE SOFTWAREHISTORY 8-130
8.7.5.27  DESCRIBE SOFTWAREUPDATE 8-130
8.7.5.28  DESCRIBE THRESHOLD 8-131
8.7.5.29  DESCRIBE USER 8-132

8.7.6  DROP
8.7.6.1  DROP ALERTHISTORY 8-133
8.7.6.2  DROP CELL 8-133
8.7.6.3  DROP CELLDISK 8-135
8.7.6.4  DROP FLASHCACHE 8-137
8.7.6.5  DROP FLASHLOG 8-138
8.7.6.6  DROP GRIDDISK 8-139
8.7.6.7  DROP PMEMCACHE 8-140
8.7.6.8  DROP PMEMLOG 8-141
8.7.6.9  DROP QUARANTINE 8-142
8.7.6.10  DROP ROLE 8-142
8.7.6.11  DROP SOFTWAREHISTORY 8-142
8.7.6.12  DROP THRESHOLD 8-143
8.7.6.13  DROP USER 8-143

8.7.7  EXIT 8-144

8.7.8  EXPORT CELLDISK 8-144

8.7.9  GRANT
8.7.9.1  GRANT PRIVILEGE 8-146
8.7.9.2  GRANT ROLE 8-147

8.7.10  HELP 8-148

8.7.11  IMPORT CELLDISK 8-148

8.7.12  LIST
8.7.12.1  LIST ACTIVEREQUEST 8-152
8.7.12.2  LIST ALERTDEFINITION 8-153
8.7.12.3  LIST ALERTHISTORY 8-153
8.7.12.4  LIST CELL 8-156
8.7.12.5  LIST CELLDISK 8-158
9  Using the dcli Utility

9.1  Overview of the dcli Utility  9-1
9.2  dcli Syntax  9-2
9.3  dcli Examples  9-4

9.3.1  Using dcli to Set up SSH User-equivalence for a Current User  9-5
### 9.3 Using dcli
- **9.3.2 Using dcli with the -n Option**  
- **9.3.3 Using dcli with the -r Option**  
- **9.3.4 Using dcli with the -v Option**  
- **9.3.5 Using dcli with the -t Option**  
- **9.3.6 Using dcli with the -f Option**  
- **9.3.7 Using dcli with the --vmstat Option**  
- **9.3.8 Using dcli with the --hidestderr Option**  
- **9.3.9 Using dcli with the --showbanner Option**  
- **9.3.10 Using dcli to Change an IORM Plan**  
- **9.3.11 Using dcli with a Script**  
- **9.3.12 Using dcli to List Grid Disk Status**  
- **9.3.13 Using dcli to List Alert History Information**  
- **9.3.14 Using dcli to List Alert History Where examinedby is Not Set**  
- **9.3.15 Using dcli to List Current Metric Alert State**  
- **9.3.16 Using dcli to List Specific Metric Current Objects in a Group**  
- **9.3.17 Using dcli to List Specific Metric Current Objects**  
- **9.3.18 Using dcli to List Physical Disks**  
- **9.3.19 Using dcli to List Cell Disks with Free Space**  
- **9.3.20 Using dcli to View Alert History**  

### 9.4 Using dcli to Change an IORM Plan  
**9.3.10**

### 9.5 Using dcli with a Script  
**9.3.11**

### 9.6 Using dcli to List Grid Disk Status  
**9.3.12**

### 9.7 Using dcli to List Alert History Information  
**9.3.13**

### 9.8 Using dcli to List Alert History Where examinedby is Not Set  
**9.3.14**

### 9.9 Using dcli to List Current Metric Alert State  
**9.3.15**

### 9.10 Using dcli to List Specific Metric Current Objects in a Group  
**9.3.16**

### 9.11 Using dcli to List Specific Metric Current Objects  
**9.3.17**

### 9.12 Using dcli to List Physical Disks  
**9.3.18**

### 9.13 Using dcli to List Cell Disks with Free Space  
**9.3.19**

### 9.14 Using dcli to View Alert History  
**9.3.20**

### 10 Setting Up Oracle Exadata Storage Snapshots

#### 10.1 Overview of Exadata Snapshots
- **10.1.1 Exadata Snapshot Support of Exadata Features**  
- **10.1.2 Separate Test/Development and Production Environments**  
- **10.1.3 Types of Exadata Snapshots**  
- **10.1.4 Hierarchical Snapshot Databases**  
- **10.1.5 Sparse Test Masters**  

#### 10.2 Prerequisites for Exadata Snapshot Databases
- **10.2.1**

#### 10.3 Exadata Snapshot Concepts
- **10.3.1 Sparse Database and Sparse Files**  
- **10.3.2 Sparse Grid Disks**  
- **10.3.3 Sparse Disk Groups**  

#### 10.4 Sparse Disk Sizing and Allocation Methodology
- **10.4.1 Sizing Steps for New Sparse Disk Groups**  

#### 10.5 Refresh Considerations, or Lifecycle of Exadata Snapshots
- **10.5.1**

#### 10.6 Using an Oracle Data Guard Standby Database as the Test Master
- **10.6.1**

#### 10.7 Managing Exadata Snapshots
- **10.7.1 Creating Sparse Grid Disks**  
- **10.7.1.1 Calculating the Physical Size for Grid Disks**  

#### 10.7.2 Creating Sparse Disk Groups
- **10.7.2.1**
B Installation Information

B.1 Installation of Oracle Exadata Storage Server B-1
   B.1.1 Installing Oracle Exadata Storage Server Hardware B-1
   B.1.2 Installing the Oracle Exadata System Software on a Cell B-2
   B.1.3 Updates Made During Installation B-3
   B.1.4 Online Documentation B-3
B.2 Background Processes in the Oracle Exadata Storage Server Environment B-3
   B.2.1 diskmon Process B-3
   B.2.2 XDMG Process B-4
   B.2.3 XDWK Process B-4

C Exadata-Specific Information in Oracle Database Dictionary Views

C.1 Oracle Database Dictionary Views C-1
   C.1.1 Using the V$CELL and GV$CELL Views to Display Oracle Exadata Storage Server Identification C-1
   C.1.2 About the V$CELL_ Views C-2
   C.1.3 Using V$BACKUP_DATAFILE with Oracle Exadata Storage Server C-3
   C.1.4 Using V$ASM_DISK_SPARSE and V$ASM_DISKGROUP_SPARSE to Monitor Sparse Disks C-3
C.2 Automatic Workload Repository Views C-4
   C.2.1 DBA_HIST_ASM_BAD_DISK C-5
   C.2.2 DBA_HIST_ASM_DISKGROUP C-5
   C.2.3 DBA_HIST_ASM_DISKGROUP_STAT C-6
   C.2.4 DBA_HIST_ASM_DISK_STAT_SUMMARY C-6
   C.2.5 DBA_HIST_CELL_CONFIG C-7
   C.2.6 DBA_HIST_CELL_CONFIG_DETAIL C-8
   C.2.7 DBA_HIST_CELL_DB C-9
   C.2.8 DBA_HIST_CELL_DISKTYPE C-11
   C.2.9 DBA_HIST_CELL_DISK_NAME C-12
   C.2.10 DBA_HIST_CELL_DISK_SUMMARY C-13
   C.2.11 DBA_HIST_CELL_GLOBAL C-15
   C.2.12 DBA_HIST_CELL_GLOBAL_SUMMARY C-16
   C.2.13 DBA_HIST_CELL_IOREASON C-17
   C.2.14 DBA_HIST_CELL_IOREASON_NAME C-17
   C.2.15 DBA_HIST_CELL_METRIC_DESC C-18
   C.2.16 DBA_HIST_CELL_NAME C-18
   C.2.17 DBA_HIST_CELL_OPEN_ALERTS C-18
Oracle Exadata System Software Accessibility Recommendations

D.1 Tips on Using Screen Readers and Braille Displays  D-1
D.2 Tips on Using Screen Magnifiers  D-2
D.3 Tips on Using Exawatcher Charts  D-3
D.4 Oracle Exadata Deployment Assistant (OEDA) Web interface Accessibility  D-3
Preface

Oracle Exadata System Software User’s Guide describes how to initialize and administer Oracle Exadata System Software. This guide describes the Oracle Exadata System Software product and its components, as well as Oracle Exadata System Software administrative and deployment procedures. This preface contains the following topics:

• Audience
• Documentation Accessibility
• Related Documents
• Conventions
• Audience
• Documentation Accessibility
• Related Documents
• Conventions

Audience

Oracle Exadata System Software User’s Guide is intended for Oracle Database and storage administrators who perform the following tasks:

• Configure Oracle Exadata System Software
• Manage Oracle Exadata System Software
• Troubleshoot Oracle Exadata System Software

Documentation Accessibility

For information about Oracle’s commitment to accessibility, visit the Oracle Accessibility Program website at http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc.

Access to Oracle Support

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info or visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs if you are hearing impaired.
Related Documents

For additional information, see the following Oracle resources:

• Oracle Exadata Database Machine System Overview
• Oracle Exadata Database Machine Installation and Configuration Guide
• Oracle Exadata Database Machine Maintenance Guide
• Oracle Engineered Systems Extending and Multi-Rack Cabling Guide
• Oracle Exadata Database Machine Security Guide
• Oracle Database 2 Day DBA
• Oracle Database Administrator's Guide
• Oracle Database Concepts
• Oracle Automatic Storage Management Administrator's Guide
• Oracle Database Error Messages Reference
• Oracle Database 2 Day + Real Application Clusters Guide
• Oracle Clusterware Administration and Deployment Guide
• Oracle Real Application Clusters Administration and Deployment Guide
• Platform-specific Oracle Database, Oracle Clusterware, and Oracle Real Application Clusters installation guides

Conventions

The following text conventions are used in this document:

<table>
<thead>
<tr>
<th>Convention</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>boldface</strong></td>
<td>Boldface type indicates graphical user interface elements associated with an action, emphasis, or terms defined in text or the glossary.</td>
</tr>
<tr>
<td><em>italic</em></td>
<td>Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.</td>
</tr>
<tr>
<td><code>monospace</code></td>
<td>Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.</td>
</tr>
<tr>
<td><code>$</code> prompt</td>
<td>The dollar sign ($) prompt indicates a command run as the <code>oracle</code> user.</td>
</tr>
<tr>
<td><code>#</code> prompt</td>
<td>The pound (#) prompt indicates a command that is run as the <code>root</code> user.</td>
</tr>
</tbody>
</table>
1
Introducing Oracle Exadata System Software

This chapter introduces Oracle Exadata System Software.

• **Overview of Oracle Exadata System Software**
  Oracle Exadata Storage Server is a highly optimized storage server that runs Oracle Exadata System Software to store and access Oracle Database data.

• **Key Features of Oracle Exadata System Software**
  This section describes the key features of Oracle Exadata System Software.

• **Oracle Exadata System Software Components**
  This section provides a summary of the following Oracle Exadata System Software components.

1.1 Overview of Oracle Exadata System Software

Oracle Exadata Storage Server is a highly optimized storage server that runs Oracle Exadata System Software to store and access Oracle Database data.

With traditional storage, data is transferred to the database server for processing. In contrast, Oracle Exadata System Software provides database-aware storage services, such as the ability to offload SQL and other database processing from the database server, while remaining transparent to the SQL processing and database applications. Oracle Exadata Database Machine storage servers process data at the storage level, and pass only what is needed to the database servers.

Oracle Exadata System Software is installed on both the storage servers and the database servers. Oracle Exadata System Software offloads some SQL processing from the database server to the storage servers. Oracle Exadata System Software enables function shipping between the database instance and the underlying storage, in addition to traditional data shipping. Function shipping greatly reduces the amount of data processing that must be done by the database server. Eliminating data transfers and database server workload can greatly benefit query processing operations that often become bandwidth constrained. Eliminating data transfers can also provide a significant benefit to online transaction processing (OLTP) systems that include large batch and report processing operations.

The hardware components of Oracle Exadata Storage Server are carefully chosen to match the needs of high performance processing. The Oracle Exadata System Software is optimized to maximize the advantage of the hardware components. Each storage server delivers outstanding processing bandwidth for data stored on disk, often several times better than traditional solutions.

Oracle Exadata Database Machine storage servers use state-of-the-art RDMA Network Fabric interconnections between servers and storage. Each RDMA Network Fabric link provides bandwidth of 40 GB for InfiniBand Network Fabric or 100 GB for RoCE Network Fabric. Additionally, the interconnection protocol uses direct data placement, also referred to as direct memory access (DMA), to ensure low CPU
overhead by directly moving data from the wire to database buffers with no extra copies. The RDMA Network Fabric has the flexibility of a LAN network with the efficiency of a storage area network (SAN). With an RDMA Network Fabric network, Oracle Exadata Database Machine eliminates network bottlenecks that could reduce performance. This RDMA Network Fabric network also provides a high performance cluster interconnection for Oracle Real Application Clusters (Oracle RAC) servers.

Each storage server has disk (rotating) storage. There are several configurations of storage server each configured to maximize some aspect of storage based on your requirements. You can specify storage servers with either Exadata Smart Flash Cache installed or Exadata Smart Flash Cache available as an option. Some storage servers also have Persistent Memory (PMEM) cache. Each of the cache configuration use RDMA providing maximum transfer rates.

The Oracle Exadata Database Machine architecture scales to any level of performance. To achieve higher performance or greater storage capacity, you add more storage servers (cells) to the configuration. As more storage servers are added, capacity and performance increase linearly. Data is mirrored across storage servers to ensure that the failure of a storage server does not cause loss of data or availability. The scale-out architecture achieves near infinite scalability, while lowering costs by allowing storage to be purchased incrementally on demand.

---

**Note:**

Oracle Exadata System Software must be used with Oracle Exadata Database Machine storage server hardware, and only supports Oracle databases on the database servers of Oracle Exadata Database Machines. Information is available on My Oracle Support at [http://support.oracle.com](http://support.oracle.com) and on the Products page of Oracle Technology Network at [http://www.oracle.com/technetwork/index.html](http://www.oracle.com/technetwork/index.html)

---

### 1.2 Key Features of Oracle Exadata System Software

This section describes the key features of Oracle Exadata System Software.

- **Reliability, Modularity, and Cost-Effectiveness**
  Oracle Exadata System Software enables cost-effective modular storage hardware to be used in a scale-out architecture while providing a high level of availability and reliability.

- **Compatibility with Oracle Database**
  When the minimum required versions are met, all Oracle Database features are fully supported with Oracle Exadata System Software.

- **Smart Flash Technology**
  The Exadata Smart Flash Cache feature of the Oracle Exadata System Software intelligently caches database objects in flash memory, replacing slow, mechanical I/O operations to disk with very rapid flash memory operations.
• Persistent Memory Accelerator and RDMA
  Persistent Memory (PMEM) Accelerator provides direct access to persistent memory using Remote Direct Memory Access (RDMA), enabling faster response times and lower read latencies.

• Centralized Storage
  You can use Oracle Exadata Storage Server to consolidate your storage requirements into a central pool that can be used by multiple databases.

• I/O Resource Management (IORM)
  I/O Resource Management (IORM) and the Oracle Database Resource Manager enable multiple databases and pluggable databases to share the same storage while ensuring that I/O resources are allocated across the various databases.

• Exadata Hybrid Columnar Compression
  Exadata Hybrid Columnar Compression stores data using column organization, which brings similar values close together and enhances compression ratios.

• In-Memory Columnar Format Support
  In an Oracle Exadata Database Machine environment, the data is automatically stored in In-Memory columnar format in the flash cache when it will improve performance.

• Offloading of Data Search and Retrieval Processing
  One of the most powerful features of Oracle Exadata System Software is that it offloads the data search and retrieval processing to the storage servers.

• Offloading of Incremental Backup Processing
  To optimize the performance of incremental backups, the database can offload block filtering to Oracle Exadata Storage Server.

• Fault Isolation with Quarantine
  Oracle Exadata System Software has the ability to learn from the past events to avoid a potential fatal error.

• Protection Against Data Corruption
  Data corruptions, while rare, can have a catastrophic effect on a database, and therefore on a business.

• Fast File Creation
  File creation operations are offloaded to Oracle Exadata Storage Servers.

• Storage Index
  Oracle Exadata Storage Servers maintain a storage index which contains a summary of the data distribution on the disk.

1.2.1 Reliability, Modularity, and Cost-Effectiveness

Oracle Exadata System Software enables cost-effective modular storage hardware to be used in a scale-out architecture while providing a high level of availability and reliability.

All single points of failure are eliminated in the Oracle Exadata Storage Server architecture by data mirroring, fault isolation technology, and protection against disk and other storage hardware failure. Even brownouts are limited or eliminated when failures occur.

In the Oracle Exadata Storage Server architecture, one or more storage cells can support one or more databases. The placement of data is transparent to database users and applications. Storage cells use Oracle Automatic Storage Management.
Oracle ASM) to distribute data evenly across the cells. Because Oracle Exadata Storage Servers support dynamic disk insertion and removal, the online dynamic data redistribution feature of Oracle ASM ensures that data is appropriately balanced across the newly added, or remaining, disks without interrupting database processing. Oracle Exadata Storage Servers also provides data protection from disk and cell failures.

1.2.2 Compatibility with Oracle Database

When the minimum required versions are met, all Oracle Database features are fully supported with Oracle Exadata System Software.

Oracle Exadata System Software works equally well with single-instance or Oracle Real Application Clusters (Oracle RAC) deployments of Oracle Database. Oracle Data Guard, Oracle Recovery Manager (RMAN), Oracle GoldenGate, and other database features are managed the same with Exadata storage cells as with traditional storage. This enables database administrators to use the same tools with which they are familiar.

Refer to My Oracle Support Doc ID 888828.1 for a complete list of the minimum required software versions.

Related Topics
• Exadata Database Machine and Exadata Storage Server Supported Versions (My Oracle Support Doc ID 888828.1)

1.2.3 Smart Flash Technology

The Exadata Smart Flash Cache feature of the Oracle Exadata System Software intelligently caches database objects in flash memory, replacing slow, mechanical I/O operations to disk with very rapid flash memory operations.

• **Exadata Smart Flash Cache**
  Exadata Smart Flash Cache holds frequently accessed data in high-performance flash storage, while most data is kept in very cost-effective disk storage.

• **Write-Back Flash Cache**
  Write-Back Flash Cache enables write I/Os directly to Exadata Smart Flash Cache.

• **Exadata Smart Flash Log**
  Exadata Smart Flash Log improves transaction response times and increases overall database throughput for I/O intensive workloads by accelerating performance-critical log write operations.

1.2.3.1 Exadata Smart Flash Cache

Exadata Smart Flash Cache holds frequently accessed data in high-performance flash storage, while most data is kept in very cost-effective disk storage.

Caching occurs automatically and requires no user or administrator effort.

Exadata Smart Flash Cache intelligently determines the data that is most useful to cache based on data usage, access patterns, and hints from the database that indicate the type of data being accessed. It also avoids caching data that will never be reused or will not fit into the cache.
Although it is generally not required or recommended, Oracle Exadata System Software also enables administrators to override the default caching policy and keep specific table and index segments in or out of the cache.

Originally, Exadata Smart Flash Cache operated exclusively in Write-Through mode. In Write-Through mode, database writes go to disk first, and subsequently populate Flash Cache. If a flash device fails with Exadata Smart Flash Cache operating in Write-Through mode, there is no data loss because the data is already on disk.

1.2.3.2 Write-Back Flash Cache

Write-Back Flash Cache enables write I/Os directly to Exadata Smart Flash Cache. With Exadata Smart Flash Cache in Write-Back mode, database writes go to Flash Cache first and later to disk. Write-Back mode was introduced with Oracle Exadata System Software release 11.2.3.2.0.

Write-intensive applications can benefit significantly from Write-Back mode by taking advantage of the fast latencies provided by flash. If your application writes intensively and if you experience high I/O latency or significant waits for free buffer waits, then you should consider using Write-Back Flash Cache.

With Exadata Smart Flash Cache in Write-Back mode, the total amount of disk I/O also reduces when the cache absorbs multiple writes to the same block before writing it to disk. The saved I/O bandwidth can be used to increase the application throughput or service other workloads.

However, if a flash device fails while using Write-Back mode, data that is not yet persistent to disk is lost and must be recovered from a mirror copy. For this reason, Write-Back mode is recommended in conjunction with high redundancy ASM disk groups.

The contents of the Write-Back Flash Cache is persisted across reboots, eliminating any warm-up time needed to populate the cache.

1.2.3.3 Exadata Smart Flash Log

Exadata Smart Flash Log improves transaction response times and increases overall database throughput for I/O intensive workloads by accelerating performance-critical log write operations.

The time to commit user transactions is very sensitive to the latency of log write operations. In addition, many performance-critical database algorithms, such as space management and index splits, are very sensitive to log write latency.

Although the disk controller has a large battery-backed DRAM cache that can accept writes very quickly, some write operations to disk can still be slow during busy periods when the disk controller cache is occasionally filled with blocks that have not been written to disk. Noticeable performance issues can arise even with relatively few slow redo log write operations.

Exadata Smart Flash Log reduces the average latency for performance-sensitive redo log write I/O operations, thereby eliminating performance bottlenecks that may occur due to slow redo log writes. Exadata Smart Flash Log removes latency spikes by simultaneously performing redo log writes to two media devices. The redo write is acknowledged as soon as the first write to either media device completes.
Prior to Oracle Exadata System Software release 20.1, Exadata Smart Flash Log performs simultaneous writes to disk and flash storage. With this configuration, Exadata Smart Flash Log improves average log write latency and increases overall database throughput. But, because all log writes must eventually persist to disk, this configuration is limited by the overall disk throughput, and provides little relief for applications that are constrained by disk throughput.

Oracle Exadata System Software release 20.1 adds a further optimization, known as Smart Flash Log Write-Back, that uses Exadata Smart Flash Cache in Write-Back mode instead of disk storage. With this configuration, Exadata Smart Flash Log improves average log write latency and overall log write throughput to eliminate logging bottlenecks for demanding throughput-intensive applications.

1.2.4 Persistent Memory Accelerator and RDMA

Persistent Memory (PMEM) Accelerator provides direct access to persistent memory using Remote Direct Memory Access (RDMA), enabling faster response times and lower read latencies.

Starting with Oracle Exadata System Software release 19.3.0, workloads that require ultra low response time, such as stock trades and IOT devices, can take advantage of PMEM and RDMA in the form of a PMEM Cache and PMEM Log.

PMEM is a new storage tier that was first released on Exadata X8M. When database clients read from the PMEM cache, the client software performs an RDMA read of the cached data, which bypasses the storage server software and delivers results much faster than Exadata Smart Flash Cache.

PMEM cache works in conjunction with Exadata Smart Flash Cache. The following table describes the supported caching mode combinations when PMEM Cache is configured:

<table>
<thead>
<tr>
<th>PMEM Cache Mode</th>
<th>Flash Cache Mode</th>
<th>Supported Configuration?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write-Through</td>
<td>Write-Through</td>
<td>Yes. This is the default configuration for High Capacity (HC) servers with Normal Redundancy.</td>
</tr>
<tr>
<td>Write-Through</td>
<td>Write-Back</td>
<td>Yes. This is the default configuration for HC servers with High Redundancy. This is also the default configuration for Extreme Flash (EF) servers.</td>
</tr>
<tr>
<td>Write-Back</td>
<td>Write-Back</td>
<td>Yes.</td>
</tr>
<tr>
<td>Write-Back</td>
<td>Write-Through</td>
<td>No. Without the backing of Write-Back Flash Cache, write-intensive workloads can overload the PMEM Cache in Write-Back mode.</td>
</tr>
</tbody>
</table>

If PMEM Cache is not configured, Exadata Smart Flash Cache is supported in both Write-Back and Write-Through modes.

Redo log writes are critical database operations and need to complete in a timely manner to prevent load spikes or stalls. Exadata Smart Flash Log is designed to prevent redo write latency outliers. PMEM Log helps to further reduce redo log write latency by using PMEM and RDMA.

With PMEM Log, database clients send redo log I/O buffers directly to PMEM on the storage servers using RDMA, thereby reducing transport latency. The cell server
then writes the redo to Exadata Smart Flash Log (if enabled) and disk at a later time.

Reduced redo log write latency improves OLTP performance, resulting in higher transaction throughput. In cases where PMEM Log is bypassed, Exadata Smart Flash Log can still be used.

1.2.5 Centralized Storage

You can use Oracle Exadata Storage Server to consolidate your storage requirements into a central pool that can be used by multiple databases.

Oracle Exadata System Software with Oracle Automatic Storage Management (Oracle ASM) evenly distributes the data and I/O load for every database across available disks in the storage pool. Every database can use all of the available disks to achieve superior I/O rates. Oracle Exadata Storage Servers can provide higher efficiency and performance at a lower cost while also lowering your storage administration overhead.

1.2.6 I/O Resource Management (IORM)

I/O Resource Management (IORM) and the Oracle Database Resource Manager enable multiple databases and pluggable databases to share the same storage while ensuring that I/O resources are allocated across the various databases.

Oracle Exadata System Software works with IORM and Oracle Database Resource Manager to ensure that customer-defined policies are met, even when multiple databases share the same set of storage servers. As a result, one database cannot monopolize the I/O bandwidth and degrade the performance of the other databases.

IORM enables storage cells to service I/O resources among multiple applications and users across all databases in accordance with sharing and prioritization levels established by the administrator. This improves the coexistence of online transaction processing (OLTP) and reporting workloads, because latency-sensitive OLTP applications can be given a larger share of disk and flash I/O bandwidth than throughput-sensitive batch applications. Oracle Database Resource Manager enables the administrator to control processor utilization on the database host on a per-application basis. Combining IORM and Oracle Database Resource Manager enables the administrator to establish more accurate policies.

IORM also manages the space utilization for Exadata Smart Flash Cache and PMEM cache. Critical OLTP workloads can be guaranteed space in Exadata Smart Flash Cache or PMEM cache to provide consistent performance.

IORM for a database or pluggable database (PDB) is implemented and managed from the Oracle Database Resource Manager. Oracle Database Resource Manager in the database instance communicates with the IORM software in the storage cell to manage user-defined service-level targets. Database resource plans are administered from the database, while interdatabase plans are administered on the storage cell.

Related Topics

- Managing I/O Resources
1.2.7 Exadata Hybrid Columnar Compression

Exadata Hybrid Columnar Compression stores data using column organization, which brings similar values close together and enhances compression ratios.

Using Exadata Hybrid Columnar Compression, data is organized into sets of rows called compression units. Within a compression unit, data is organized by column and then compressed. Each row is self-contained within a compression unit.

Database operations work transparently against compressed objects, so no application changes are required. The database compresses data manipulated by any SQL operation, although compression levels are higher for direct path loads.

You can specify the following types of Exadata Hybrid Columnar Compression, depending on your requirements:

- Warehouse compression: This type of compression is optimized for query performance, and is intended for data warehouse applications.
- Archive compression: This type of compression is optimized for maximum compression levels, and is intended for historic data and data that does not change.

Assume that you apply Exadata Hybrid Columnar Compression to a `daily_sales` table. At the end of every day, the table is populated with items and the number sold, with the item ID and date forming a composite primary key. A row subset is shown in the following table.

### Table 1-1  Sample Table `daily_sales`

<table>
<thead>
<tr>
<th>Item_ID</th>
<th>Date</th>
<th>Num_Sold</th>
<th>Shipped_From</th>
<th>Restock</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>01-JUN-07</td>
<td>2</td>
<td>WAREHOUSE1</td>
<td>Y</td>
</tr>
<tr>
<td>1001</td>
<td>01-JUN-07</td>
<td>0</td>
<td>WAREHOUSE3</td>
<td>N</td>
</tr>
<tr>
<td>1002</td>
<td>01-JUN-07</td>
<td>1</td>
<td>WAREHOUSE3</td>
<td>N</td>
</tr>
<tr>
<td>1003</td>
<td>01-JUN-07</td>
<td>0</td>
<td>WAREHOUSE2</td>
<td>N</td>
</tr>
<tr>
<td>1004</td>
<td>01-JUN-07</td>
<td>2</td>
<td>WAREHOUSE1</td>
<td>N</td>
</tr>
<tr>
<td>1005</td>
<td>01-JUN-07</td>
<td>1</td>
<td>WAREHOUSE2</td>
<td>N</td>
</tr>
</tbody>
</table>

The database stores a set of rows in an internal structure called a compression unit. For example, assume that the rows in the previous table are stored in one unit. Exadata Hybrid Columnar Compression stores each unique value from column 4 with metadata that maps the values to the rows. Conceptually, the compressed value can be represented as:

WAREHOUSE1WAREHOUSE3WAREHOUSE2

The database then compresses the repeated word `WAREHOUSE` in this value by storing it once and replacing each occurrence with a reference. If the reference is smaller than the original word, then the database achieves compression. The compression benefit is particularly evident for the `Date` column, which contains only one unique value.

As shown in the following illustration, each compression unit can span multiple data blocks. The values for a particular column may or may not span multiple blocks.
Exadata Hybrid Columnar Compression has implications for row locking. When an update occurs for a row in an uncompressed data block, only the updated row is locked. In contrast, the database must lock all rows in the compression unit if an update is made to any row in the unit. Updates to rows using Exadata Hybrid Columnar Compression cause rowids to change.

Note:

When tables use Exadata Hybrid Columnar Compression, Oracle DML locks larger blocks of data (compression units) which may reduce concurrency.

Oracle Database supports four methods of table compression.

Table 1-2 Table Compression Methods

<table>
<thead>
<tr>
<th>Table Compression Method</th>
<th>Compression Level</th>
<th>CPU Overhead</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic compression</td>
<td>High</td>
<td>Minimal</td>
<td>DSS</td>
</tr>
<tr>
<td>OLTP compression</td>
<td>High</td>
<td>Minimal</td>
<td>OLTP, DSS</td>
</tr>
<tr>
<td>Warehouse compression</td>
<td>Higher (compression level depends on compression level specified (LOW or HIGH))</td>
<td>Higher (CPU overhead depends on compression level specified (LOW or HIGH))</td>
<td>DSS</td>
</tr>
<tr>
<td>Archive compression</td>
<td>Highest (compression level depends on compression level specified (LOW or HIGH))</td>
<td>Highest (CPU overhead depends on compression level specified (LOW or HIGH))</td>
<td>Archiving</td>
</tr>
</tbody>
</table>

Warehouse compression and archive compression achieve the highest compression levels because they use Exadata Hybrid Columnar Compression technology. Exadata Hybrid Columnar Compression technology uses a modified form of columnar storage instead of row-major storage. This enables the database to store similar data together,
which improves the effectiveness of compression algorithms. Because Exadata Hybrid Columnar Compression requires high CPU overhead for DML, use it only for data that is updated infrequently.

The higher compression levels of Exadata Hybrid Columnar Compression are achieved only with data that is direct-path inserted. Conventional inserts and updates are supported, but result in a less compressed format, and reduced compression level.

The following table lists characteristics of each table compression method.

**Table 1-3  Table Compression Characteristics**

<table>
<thead>
<tr>
<th>Table Compression Method</th>
<th>CREATE/ALTER TABLE Syntax</th>
<th>Direct-Path Insert</th>
<th>DML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic compression</td>
<td>COMPRESS [BASIC]</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>COMPRESS and COMPRESS BASIC are equivalent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLTP compression</td>
<td>COMPRESS FOR OLTP</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Warehouse compression</td>
<td>COMPRESS FOR QUERY [LOW</td>
<td>HIGH]</td>
<td>Yes</td>
</tr>
<tr>
<td>Archive compression</td>
<td>COMPRESS FOR ARCHIVE [LOW</td>
<td>HIGH]</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The COMPRESS FOR QUERY HIGH option is the default data warehouse compression mode. It provides good compression and performance. The COMPRESS FOR QUERY LOW option should be used in environments where load performance is critical. It loads faster than data compressed with the COMPRESS FOR QUERY HIGH option.

The COMPRESS FOR ARCHIVE LOW option is the default archive compression mode. It provides a high compression level and good query performance. It is ideal for infrequently-accessed data. The COMPRESS FOR ARCHIVE HIGH option should be used for data that is rarely accessed.

A compression advisor, provided by the DBMS_COMPRESSION package, helps you determine the expected compression level for a particular table with a particular compression method.

You specify table compression with the COMPRESS clause of the CREATE TABLE command. You can enable compression for an existing table by using these clauses in an ALTER TABLE statement. In this case, only data that is inserted or
updated is compressed after compression is enabled. Similarly, you can disable table compression for an existing compressed table with the `ALTER TABLE...NOCOMPRESS` command. In this case, all data that was already compressed remains compressed, and new data is inserted uncompressed.

1.2.8 In-Memory Columnar Format Support

In an Oracle Exadata Database Machine environment, the data is automatically stored in In-Memory columnar format in the flash cache when it will improve performance.

Oracle Exadata Database Machine supports all of the In-Memory optimizations, such as accessing only the compressed columns required, SIMD vector processing, storage indexes, and so on.

If you set the `INMEMORY_SIZE` database initialization parameter to a non-zero value (requires the Oracle Database In-Memory option), then objects accessed using a Smart Scan are brought into the flash cache and are automatically converted into the In-Memory columnar format. The data is converted initially into a columnar cache format, which is different from Oracle Database In-Memory’s columnar format. The data is rewritten in the background into Oracle Database In-Memory columnar format. As a result, all subsequent accesses to the data benefit from all of the In-Memory optimizations when that data is retrieved from the flash cache.

Any write to an in-memory table does not invalidate the entire columnar cache of that table. It only invalidates the columnar cache unit of the disk region in which the block resides. For subsequent scans after a table update, a large part of the table is still in the columnar cache. The scans can still make use of the columnar cache, except for the units in which the writes were made. For those units, the query uses the original block version to get the data. After a sufficient number of scans, the invalidated columnar cache units are automatically repopulated in the columnar format.

A new segment-level attribute, `CELLMEMORY`, has also been introduced to help control which objects should not be populated into flash using the In-Memory columnar format and which type of compression should be used. Just like the `INMEMORY` attribute, you
can specify different compression levels as sub-clauses to the \textit{CELLMEMORY} attribute. However, not all of the \textit{INMEMORY} compression levels are available; only \texttt{MEMCOMPRESS FOR QUERY LOW} and \texttt{MEMCOMPRESS FOR CAPACITY LOW} (default). You specify the \textit{CELLMEMORY} attribute using a SQL command, such as the following:

\texttt{ALTER TABLE trades \textit{CELLMEMORY MEMCOMPRESS FOR QUERY LOW}}

The \texttt{PRIORTY} sub-clause available with Oracle Database In-Memory is not available on Oracle Exadata Database Machine because the process of populating the flash cache on Exadata storage servers is different from populating DRAM in the In-Memory column store on Oracle Database servers.

1.2.9 Offloading of Data Search and Retrieval Processing

One of the most powerful features of Oracle Exadata System Software is that it offloads the data search and retrieval processing to the storage servers.

One of the main advantages of Exadata Smart Scan Offload, or simply Smart Scan is that it uses storage server CPU, freeing database server from the IO and includes decompression of data stored in compressed format. Smart Scan further reduces the processing on the database server by performing predicate filtering, which entails evaluating database predicates to optimize the performance of certain classes of bulk data processing.

Oracle Database can optimize the performance of queries that perform full table, fast full index, and full bitmap index scans to evaluate selective predicates in Oracle Exadata Storage Server. The database can complete these queries faster by pushing the database expression evaluations to the storage cell. These expressions include simple SQL command predicates, such as \texttt{amount > 200}, and column projections, such as \texttt{SELECT customer\_name}. For example:

\texttt{SQL> SELECT customer\_name FROM calls WHERE amount > 200;}

In the preceding example, only rows satisfying the predicate, and specified columns are returned to the database server, eliminating unproductive data transfer to the database server.

Oracle Exadata System Software uses storage-side predicate evaluation that transfers simplified, predicate evaluation operations for table and index scans to the storage cell. This brings the table scan closer to the disk to enable a higher bandwidth, and prevents sending unmatched rows to hosts.
1.2.10 Offloading of Incremental Backup Processing

To optimize the performance of incremental backups, the database can offload block filtering to Oracle Exadata Storage Server.

This optimization is only possible when taking backups using Oracle Recovery Manager (RMAN). The offload processing is done transparently without user intervention. During offload processing, Oracle Exadata System Software filters out the blocks that are not required for the incremental backup in progress. Therefore, only the blocks that are required for the backup are sent to the database, making backups significantly faster.

**Related Topics**
- Making Incremental Backups: Quick Start

1.2.11 Fault Isolation with Quarantine

Oracle Exadata System Software has the ability to learn from the past events to avoid a potential fatal error.

When a faulty SQL statement caused a crash of the server in the past, Oracle Exadata System Software quarantines the SQL statement so that when the faulty SQL statement occurs again, Oracle Exadata System Software does not allow the SQL statement to perform Smart Scan. This reduces the chance of server software
crashes, and improves storage availability. The following types of quarantine are available:

- **SQL Plan**: Created when Oracle Exadata System Software crashes while performing Smart Scan for a SQL statement. As a result, the SQL Plan for the SQL statement is quarantined, and Smart Scan is disabled for the SQL plan.

- **Disk Region**: Created when Oracle Exadata System Software crashes while performing Smart Scan of a disk region. As a result, the 1 MB disk region being scanned is quarantined and Smart Scan is disabled for the disk region.

- **Database**: Created when Oracle Exadata System Software detects that a particular database causes instability to a cell. Instability detection is based on the number of SQL Plan Quarantines for a database. Smart Scan is disabled for the database.

- **Cell Offload**: Created when Oracle Exadata System Software detects some offload feature has caused instability to a cell. Instability detection is based on the number of Database Quarantines for a cell. Smart Scan is disabled for all databases.

- **Intra-Database Plan**: Created when Oracle Exadata System Software crashes while processing an intra-database resource plan. Consequently, the intra-database resource plan is quarantined and not enforced. Other intra-database resource plans in the same database are still enforced. Intra-database resource plans in other databases are not affected.

- **Inter-Database Plan**: Created when Oracle Exadata System Software crashes while processing an inter-database resource plan. Consequently, the inter-database resource plan is quarantined and not enforced. Other inter-database resource plans are still enforced.

- **I/O Resource Management (IORM)**: Created when Oracle Exadata System Software crashes in the I/O processing code path. IORM is effectively disabled by setting the IORM objective to basic and all resource plans are ignored.

- **Cell-to-Cell Offload**: See “Quarantine Manager Support for Cell-to-Cell Offload Operations”.

When a quarantine is created, alerts notify administrators of what was quarantined, why the quarantine was created, when and how the quarantine can be dropped manually, and when the quarantine is dropped automatically. All quarantines are automatically removed when a cell is patched or upgraded.

**CellCLI commands** are used to manually manipulate quarantines. For instance, the administrator can manually create a quarantine, drop a quarantine, change attributes of a quarantine, and list quarantines.

- **Quarantine Manager Support for Cell-to-Cell Offload Operations**

**Related Topics**

- **Using the CellCLI Utility**
  You use the Cell Control Command-Line Interface (CellCLI) utility to manage Oracle Exadata System Software.

- **ADR Alert Messages**

- **Quarantine Manager Support for Cell-to-Cell Offload Operations**

- **Oracle Database Administrator’s Guide**
1.2.11.1 Quarantine Manager Support for Cell-to-Cell Offload Operations

Minimum Exadata software required: 12.2.1.1.0

Quarantine manager support is enabled for rebalance and high throughput writes in cell-to-cell offload operations. If Exadata detects a crash during these operations, the offending operation will be quarantined, and Exadata will fall back to using non-offloaded operations.

These types of quarantines are most likely caused by incompatible versions of CELLSRV. If such quarantines occur on your system, contact Oracle Support Services.

For rebalance operations, the quarantine is based on the ASM cluster ID. Rebalance will continue using the fallback path, which is slower.

For high throughput writes that originated from a database, the quarantine is based on a combination of ASM cluster ID and database ID.

For high throughput writes that originated from a CDB or PDB, the quarantine is based on a combination of ASM cluster ID and container database ID.

To identify these types of quarantine, run the LIST QUARANTINE DETAIL command and check the value of the quarantineType attribute. Values for this attribute for these quarantines are ASM_OFFLOAD_REBALANCE and HIGH_THROUGHPUT_WRITE. For the HIGH_THROUGHPUT_WRITE type there is a database case and a CDB case.

The LIST QUARANTINE statement produces output that looks like the following:

For rebalance:

CellCLI> list quarantine detail
name: 2
asmClusterId: b6063030c0ffef8dffcc99bd18b91a62
cellsrvChecksum: 9f98483ef351a1352d567ebb1ca8aeab
clientPID: 10308
comment: None
crashReason: ORA-600[CacheGet::process:C2C_OFFLOAD_CACHEGET_CRASH]
creationTime: 2016-06-23T22:33:30-07:00
dbUniqueID: 0
dbUniqueName: UnknowndbName
incidentID: 1
quarantineMode: "FULL Quarantine"
quarantinePlan: SYSTEM
quarantineReason: Crash
quarantineType: ASM_OFFLOAD_REBALANCE
remoteHostName: slc10vwt
rpmVersion: OSS_MAIN_LINUX.X64_160623

For high throughput writes that originated from database:

CellCLI> list quarantine detail
name: 10
asmClusterId: b6063030c0ffef8dffcc99bd18b91a62

For high throughput writes that originated from the CDB (differences noted in bold):

```
CellCLI> list quarantine detail
name:                   10
asmClusterId:           eff096e82317ff87bfb2ee163731f7f7
cellsrvChecksum:        9f98483ef351a1352d567ebb1ca8aeab
clientPID:              17206
comment:                None
crashReason:            ORA-600[CacheGet::process:C2C_OFFLOAD_CACHEGET_CRASH]
creationTime:           2016-06-24T12:59:06-07:00
conDbUniqueID:          4263312973
conDbUniqueName:        WRITES
dbUniqueID:             0
dbUniqueName:           UnknownDBName
incidentID:             25
quarantineMode:         "FULL Quarantine"
quarantinePlan:         SYSTEM
quarantineReason:       Crash
quarantineType:         HIGH_THROUGHPUT_WRITE
remoteHostName:         slc10vwt
rpmVersion:             OSS_MAIN_LINUX.X64_160623
```

### 1.2.12 Protection Against Data Corruption

Data corruptions, while rare, can have a catastrophic effect on a database, and therefore on a business.

Oracle Exadata System Software takes data protection to the next level by protecting business data, not just the physical bits.

The key approach to detecting and preventing corrupted data is block checking in which the storage subsystem validates the Oracle block contents. Oracle Database validates and adds protection information to the database blocks, while Oracle Exadata System Software detects corruptions introduced into the I/O path between the database and storage. The Storage Server stops corrupted data from being written to
disk. This eliminates a large class of failures that the database industry had previously been unable to prevent.

Unlike other implementations of corruption checking, checks with Oracle Exadata System Software operate completely transparently. No parameters need to be set at the database or storage tier. These checks transparently handle all cases, including Oracle Automatic Storage Management (Oracle ASM) disk rebalance operations and disk failures.

1.2.13 Fast File Creation

File creation operations are offloaded to Oracle Exadata Storage Servers.

Operations such as CREATE TABLESPACE, which can create one or more files, have a significant increase in speed due to file creation offload.

File resize operations are also offloaded to the storage servers, which are important for auto-extensible files.

1.2.14 Storage Index

Oracle Exadata Storage Servers maintain a storage index which contains a summary of the data distribution on the disk.

The storage index is maintained automatically, and is transparent to Oracle Database. It is a collection of in-memory region indexes, prior to Exadata 12.2.1.1.0 each region index stores summaries for up to eight columns, and from Exadata 12.2.1.1.0, each region index may store summaries for up to 24 columns. If set summaries are used, the maximum number of 24 may not be achieved. There is one region index for each 1 MB of disk space. Storage indexes work with any non-linguistic data type, and work with linguistic data types similar to non-linguistic indexes.

Each region index maintains the minimum and maximum values of the columns of the table. The minimum and maximum values are used to eliminate unnecessary I/O, also known as I/O filtering. The Cell physical IO bytes saved by storage index statistic, available in the V$SYS_STAT and V$SESSTAT views, shows the number of bytes of I/O saved using storage index. The content stored in one region index is independent of the other region indexes. This makes them highly scalable, and avoids latch contention.

Queries using the following comparisons are improved by the storage index:

- Equality (=)
- Inequality (<, !=, or >)
- Less than or equal (<=)
- Greater than or equal (>=)
- IS NULL
- IS NOT NULL

Oracle Exadata System Software automatically builds Storage indexes after a query with a comparison predicate that is greater than the maximum or less than the minimum value for the column in a region, and would have benefited if a storage index had been present. Oracle Exadata System Software automatically learns which
storage indexes would have benefited a query, and then creates the storage index automatically so that subsequent similar queries benefit.

In Oracle Exadata System Software release 12.2.1.1.0 and later, when data has been stored using the in-memory format columnar cache, Oracle Exadata Database Machine stores these columns compressed using dictionary encoding. For columns with fewer than 200 distinct values, the storage index creates a very compact in-memory representation of the dictionary and uses this compact representation to filter disk reads based on equality predicates. This feature is called set membership. A more limited filtering ability extends up to 400 distinct values.

For example, suppose a region of disk holds a list of customers in the United States and Canada. When you run a query looking for customers in Mexico, Oracle Exadata Storage Server can use the new set membership capability to improve the performance of the query by filtering out disk regions that do not contain customers from Mexico. In Oracle Exadata System Software releases earlier than 12.2.1.1.0, which do not have the set membership capability, a regular storage index would be unable to filter those disk regions.

Note:
The effectiveness of storage indexes can be improved by ordering the rows based on columns that frequently appear in WHERE query clauses.

Note:
The storage index is maintained during write operations to uncompressed blocks and OLTP compressed blocks. Write operations to Exadata Hybrid Columnar Compression compressed blocks or encrypted tablespaces invalidate a region index, and only the storage index on a specific region. The storage index for Exadata Hybrid Columnar Compression is rebuilt on subsequent scans.

Example 1-1 Elimination of Disk I/O with Storage Index

The following figure shows a table and region indexes. The values in the table range from one to eight. One region index stores the minimum 1, and the maximum of 5. The other region index stores the minimum of 3, and the maximum of 8.

<table>
<thead>
<tr>
<th>Table</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Min B = 1, Max B = 5</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Min B = 3, Max B = 8</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

I/O eliminated by using storage index
For a query such as `SELECT * FROM TABLE WHERE B<2`, only the first set of rows match. Disk I/O is eliminated because the minimum and maximum of the second set of rows do not match the `WHERE` clause of the query.

**Example 1-2  Partition Pruning-like Benefits with Storage Index**

In the following figure, there is a table named `Orders` with the columns `Order_Number`, `Order_Date`, `Ship_Date`, and `Order_Item`. The table is range partitioned by `Order_Date` column.

<table>
<thead>
<tr>
<th>Order#</th>
<th>Order_Date</th>
<th>Ship_Date</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2010</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2011</td>
<td>2011</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2012</td>
<td>2012</td>
<td></td>
</tr>
</tbody>
</table>

The following query looks for orders placed since January 1, 2015:

```
SELECT count(*) FROM Orders WHERE Order_Date >= to_date ('2015-01-01', 'YYY-MM-DD')
```

Because the table is partitioned on the `Order_Date` column, the preceding query avoids scanning unnecessary partitions of the table. Queries on `Ship_Date` do not benefit from `Order_Date` partitioning, but `Ship_Date` and `Order_Number` are highly correlated with `Order_Date`. Storage indexes take advantage of ordering created by partitioning or sorted loading, and can use it with the other columns in the table. This provides partition pruning-like performance for queries on the `Ship_Date` and `Order_Number` columns.

**Example 1-3  Improved Join Performance Using Storage Index**

Using storage index allows table joins to skip unnecessary I/O operations. For example, the following query would perform an I/O operation and apply a Bloom filter to only the first block of the fact table.

```
SELECT count(*) FROM fact, dim WHERE fact.m=dim.m AND dim.product="Hard drive"
```

![Diagram showing improved join performance using storage index]
The I/O for the second block of the fact table is completely eliminated by storage index as its minimum/maximum range (5,8) is not present in the Bloom filter.

1.3 Oracle Exadata System Software Components

This section provides a summary of the following Oracle Exadata System Software components.

- **About Oracle Exadata System Software**
  Unique software algorithms in Oracle Exadata System Software implement database intelligence in storage, PCI-based flash, and RDMA Network Fabric networking to deliver higher performance and capacity at lower costs than other platforms.

- **About Oracle Automatic Storage Management**
  Oracle Automatic Storage Management (Oracle ASM) is the cluster volume manager and file system used to manage Oracle Exadata Storage Server resources.

- **About Grid RAID**
  A grid Redundant Array of Independent Disks (RAID) configuration uses Oracle ASM mirroring capabilities.

- **About Storage Server Security**
  Security for Exadata Storage Servers is enforced by identifying which clients can access storage servers and grid disks.

- **About iDB Protocol**
  The iDB protocol is a unique Oracle data transfer protocol that serves as the communications protocol among Oracle ASM, database instances, and storage cells.

- **About Oracle Exadata System Software Processes**
  Oracle Exadata System Software includes the following software processes:

  - **About Cell Management**
    Each cell in the Oracle Exadata Storage Server grid is individually managed with Cell Control Command-Line Interface (CellCLI).

  - **About Database Server Software**
    Oracle software is installed on the Exadata database servers.

  - **About Oracle Enterprise Manager for Oracle Exadata Database Machine**
    Oracle Enterprise Manager provides a complete target that enables you to monitor Oracle Exadata Database Machine, including configuration and performance, in a graphical user interface (GUI).

1.3.1 About Oracle Exadata System Software

Unique software algorithms in Oracle Exadata System Software implement database intelligence in storage, PCI-based flash, and RDMA Network Fabric networking to deliver higher performance and capacity at lower costs than other platforms.

Oracle Exadata Storage Server is a network-accessible storage device with Oracle Exadata System Software installed on it. The software communicates with the database using a specialized iDB protocol, and provides both simple I/O functionality, such as block-oriented reads and writes, and advanced I/O functionality, including predicate offload and I/O Resource Management (IORM). Each storage server has
physical disks. The physical disk is an actual device within the storage server that constitutes a single disk drive spindle.

Within the storage servers, a logical unit number (LUN) defines a logical storage resource from which a single cell disk can be created. The LUN refers to the access point for storage resources presented by the underlying hardware to the upper software layers. The precise attributes of a LUN are configuration-specific. For example, a LUN could be striped, mirrored, or both striped and mirrored.

A cell disk is an Oracle Exadata System Software abstraction built on the top of a LUN. After a cell disk is created from the LUN, it is managed by Oracle Exadata System Software and can be further subdivided into grid disks, which are directly exposed to the database and Oracle Automatic Storage Management (Oracle ASM) instances. Each grid disk is a potentially non-contiguous partition of the cell disk that is directly exposed to Oracle ASM to be used for the Oracle ASM disk group creations and expansions.

This level of virtualization enables multiple Oracle ASM clusters and multiple databases to share the same physical disk. This sharing provides optimal use of disk capacity and bandwidth. Various metrics and statistics collected on the cell disk level enable you to evaluate the performance and capacity of storage servers. IORM schedules the cell disk access in accordance with user-defined policies.

The following image illustrates how the components of a storage server (also called a cell) are related to grid disks.

- A LUN is created from a physical disk.
- A cell disk is created on a LUN. A segment of cell disk storage is used by the Oracle Exadata System Software system, referred to as the cell system area.
- Multiple grid disks can be created on a cell disk.

Figure 1-3  Oracle Exadata Storage Server Components

The following image illustrates software components in the Oracle Exadata Storage Server environment.
The figure illustrates the following environment:

- Single-instance or Oracle RAC databases access storage servers using the iDB protocol over a RDMA Network Fabric network. Each database server runs the Oracle Database and Oracle Grid Infrastructure software. Resources are managed for each database instance by Oracle Database Resource Manager (shown as DBRM).

- The database servers include Oracle Exadata System Software functionality, such as a Management Server (MS), command-line interface (DBMCLI) and OS based ExaWatcher.

- Storage servers contain cell-based utilities and processes from Oracle Exadata System Software, including:
  - Cell Server (CELLSRV)—the primary component of the Oracle Exadata System Software running in the storage server, which provides the majority of the storage server services. CELLSRV services database requests for disk I/O and provides the advanced SQL offload capabilities. CELLSRV implements the I/O Resource Management (IORM) functionality to meter out I/O bandwidth to the various databases and consumer groups issuing I/O calls on the storage server.
  - Offload Server (CELLOFLSRV <version>) — Is a helper process to the Cell Server that processes offload requests from a specific Database version. These processes allow the Storage server to respond to requests from multiple database versions residing on the same or multiple Database servers.
Management Server (MS)—the primary interface to administer, manage and query the status of the storage server. It works in cooperation with the Cell Control Command-Line Interface (CellCLI) and processes most of the commands from CellCLI.

Restart Server (RS)—monitors the heartbeat with the MS and the CELLSRV processes, and restarts the servers if they fail to respond within the allowable heartbeat period.

- Storage cells are configured on the network, and are managed by the Oracle Exadata System Software CellCLI utility.
- Each storage server contains multiple disks which store the data for the database instances on the database servers. The data is stored in disks managed by Oracle ASM.

### 1.3.2 About Oracle Automatic Storage Management

Oracle Automatic Storage Management (Oracle ASM) is the cluster volume manager and file system used to manage Oracle Exadata Storage Server resources.

Oracle ASM provides enhanced storage management by:

- Striping database files evenly across all available storage cells and disks for optimal performance.
- Using mirroring and failure groups to avoid any single point of failure.
- Enabling dynamic add and drop capability for non-intrusive cell and disk allocation, deallocation, and reallocation.
- Enabling multiple databases to share storage cells and disks.

The following topics provide a brief overview of Oracle ASM:

- **Oracle ASM Disk Groups**
  An Oracle Automatic Storage Management (Oracle ASM) disk group is the primary storage abstraction within Oracle ASM, and is composed of one or more grid disks.

- **Oracle ASM Failure Group**
  An Oracle ASM failure group is a subset of disks in an Oracle ASM disk group that can fail together because they share the same hardware.

- **Maximum Availability with Oracle ASM**
  Oracle recommends high redundancy Oracle ASM disk groups, and file placement configuration which can be automatically deployed using Oracle Exadata Deployment Assistant.

**Related Topics**
- Overview of Oracle Automatic Storage Management

### 1.3.2.1 Oracle ASM Disk Groups

An Oracle Automatic Storage Management (Oracle ASM) disk group is the primary storage abstraction within Oracle ASM, and is composed of one or more grid disks.

Oracle Exadata Storage Server grid disks appear to Oracle ASM as individual disks available for membership in Oracle ASM disk groups. Whenever possible, grid disk
names should correspond closely with Oracle ASM disk group names to assist in problem diagnosis between Oracle ASM and Oracle Exadata System Software.

Typically, Oracle Exadata Database Machine is configured with the following Oracle ASM disk groups:

- **DATA** is the primary data disk group.
- **RECO** is the primary recovery disk group, which contains the Oracle Database Fast Recovery Area (FRA).
- **SPARSE** is an optionally configured sparse disk group that is required to support Exadata snapshots.
- **XTND** is the default name for the disk group that is used to collect storage from Exadata XT (Extended) storage servers.
- **DBFS** is the system disk group that is typically configured on systems prior to Exadata X7. The DBFS disk group is primarily used to store the shared Oracle Clusterware files (Oracle Cluster Registry and voting disks) and provide some space to host Oracle Database File System (DBFS). This disk group is not configured on Exadata X7, and later, systems.

To take advantage of Oracle Exadata System Software features, such as predicate processing offload, the disk groups must contain only Oracle Exadata Storage Server grid disks, the tables must be fully inside these disk groups, and the group should have `cell.smart_scan_capable` attribute set to TRUE.

**Note:**

The Oracle Database and Oracle Grid Infrastructure software must be release 12.1.0.2.0 BP3 or later when using sparse grid disks.

**Related Topics**

- *Oracle Automatic Storage Management Administrator's Guide*

### 1.3.2.2 Oracle ASM Failure Group

An Oracle ASM failure group is a subset of disks in an Oracle ASM disk group that can fail together because they share the same hardware.

Oracle ASM considers failure groups when making redundancy decisions.

For Oracle Exadata Storage Servers, all grid disks, which consist of the Oracle ASM disk group members and candidates, can effectively fail together if the storage cell fails. Because of this scenario, all Oracle ASM grid disks sourced from a given storage cell should be assigned to a single failure group representing the cell.

For example, if all grid disks from two storage cells, A and B, are added to a single Oracle ASM disk group with normal redundancy, then all grid disks on storage cell A are designated as one failure group, and all grid disks on storage cell B are designated as another failure group. This enables Oracle Exadata System Software and Oracle ASM to tolerate the failure of either storage cell.
Failure groups for Oracle Exadata Storage Server grid disks are set by default so that the disks on a single cell are in the same failure group, making correct failure group configuration simple for Oracle Exadata Storage Servers.

You can define the redundancy level for an Oracle ASM disk group when creating a disk group. An Oracle ASM disk group can be specified with normal or high redundancy. Normal redundancy double mirrors the extents, and high redundancy triple mirrors the extents. Oracle ASM normal redundancy tolerates the failure of a single cell or any set of disks in a single cell. Oracle ASM high redundancy tolerates the failure of two cells or any set of disks in two cells. Base your redundancy setting on your desired protection level. When choosing the redundancy level, ensure the post-failure I/O capacity is sufficient to meet the redundancy requirements and performance service levels. Oracle recommends using three cells for normal redundancy. This ensures the ability to restore full redundancy after cell failure. Consider the following:

- If a cell or disk fails, then Oracle ASM automatically redistributes the cell or disk contents across the remaining disks in the disk group as long as there is enough space to hold the data. For an existing disk group using Oracle ASM redundancy, the `USABLE_FILE_MB` and `REQUIRED_FREE_MIRROR_MB` columns in the `V$ASM_DISKGROUP` view give the amount of usable space and space for redundancy, respectively.
- If a cell or disk fails, then the remaining disks should be able to generate the IOPS necessary to sustain the performance service level agreement.

After a disk group is created, the redundancy level of the disk group cannot be changed. To change the redundancy of a disk group, you must create another disk group with the appropriate redundancy, and then move the files.

Each Exadata Cell is a failure group. A normal redundancy disk group must contain at least two failure groups. Oracle ASM automatically stores two copies of the file extents, with the mirrored extents placed in different failure groups. A high redundancy disk group must contain at least three failure groups. Oracle ASM automatically stores three copies of the file extents, with each file extent in separate failure groups.

System reliability can diminish if your environment has an insufficient number of failure groups. A small number of failure groups, or failure groups of uneven capacity, can lead to allocation problems that prevent full use of all available storage.

### Related Topics
- [Administering Oracle ASM Disk Groups Using Oracle Exadata Storage Servers](#)

#### 1.3.2.3 Maximum Availability with Oracle ASM

Oracle recommends high redundancy Oracle ASM disk groups, and file placement configuration which can be automatically deployed using Oracle Exadata Deployment Assistant.

High redundancy can be configured for DATA, RECO, or any other Oracle ASM disk group with a minimum of 3 storage cells. Starting with Exadata Software release 12.1.2.3.0, the voting disks can reside in a high redundancy disk group, and additional quorum disks (essentially equivalent to voting disks) can reside on database servers if there are fewer than 5 Exadata storage cells.

Maximum availability architecture (MAA) best practice uses two main Oracle ASM disk groups: DATA and RECO. The disk groups are organized as follows:
The disk groups are striped across all disks and Oracle Exadata Storage Servers to maximize I/O bandwidth and performance, and to simplify management.

The DATA and RECO disk groups are configured for high (3-way) redundancy. The benefits of high redundancy disk groups are illustrated by the following outage scenarios:

- Double partner disk failure: Protection against loss of the database and Oracle ASM disk group due to a disk failure followed by a second partner disk failure.
- Disk failure when Oracle Exadata Storage Server is offline: Protection against loss of the database and Oracle ASM disk group when a storage server is offline and one of the storage server’s partner disks fails. The storage server may be offline because of Exadata storage planned maintenance, such as Exadata rolling storage server patching.
- Disk failure followed by disk sector corruption: Protection against data loss and I/O errors when latent disk sector corruptions exist and a partner storage disk is unavailable either due to planned maintenance or disk failure.

If the voting disk resides in a high redundancy disk group that is part of the default Exadata high redundancy deployment, the cluster and database will remain available for the above failure scenarios. If the voting disk resides on a normal redundancy disk group, then the database cluster will fail and the database has to be restarted. You can eliminate that risk by moving the voting disks to a high redundancy disk group and creating additional quorum disks on database servers.

Oracle recommends high redundancy for ALL (DATA and RECO) disk groups because it provides maximum application availability against storage failures and operational simplicity during a storage outage. In contrast, if all disk groups were configured with normal redundancy and two partner disks fail, all clusters and databases on Exadata will fail and you will lose all your data (normal redundancy does not survive double partner disk failures). Other than better storage protection, the major difference between high redundancy and normal redundancy is the amount of usable storage and write I/Os. High redundancy requires more space, and has three write I/Os instead of two. The additional write I/O normally has negligible impact with Exadata smart write-back flash cache.

The following table describes that redundancy option, as well as others, and the relative availability trade-offs. The table assumes that voting disks reside in high redundancy disk group. Refer to *Oracle Exadata Database Machine Maintenance Guide* to migrate voting disks to high redundancy disk group for existing high redundancy disk group configurations.

<table>
<thead>
<tr>
<th>Redundancy Option</th>
<th>Availability Implications</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High redundancy for ALL (DATA and RECO)</td>
<td>Zero application downtime and zero data loss for the preceding storage outage scenarios if voting disks reside in high redundancy disk group. If voting disks currently reside in normal redundancy disk group, refer to <em>Oracle Exadata Database Machine Maintenance Guide</em> to migrate them to high redundancy disk group.</td>
<td>Use this option for best storage protection and operational simplicity for mission-critical applications. Requires more space for higher redundancy.</td>
</tr>
<tr>
<td>Redundancy Option</td>
<td>Availability Implications</td>
<td>Recommendation</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>High redundancy for DATA only</strong></td>
<td>Zero application downtime and zero data loss for preceding storage outage scenarios. This option requires an alternative archive destination.</td>
<td>Use this option for best storage protection for DATA with slightly higher operational complexity. More available space than high redundancy for ALL. Refer to My Oracle Support note 2059780.1 for details.</td>
</tr>
</tbody>
</table>
| **High redundancy for RECO only** | Zero data loss for the preceding storage outage scenarios.                               | Use this option when longer recovery times are acceptable for the preceding storage outage scenarios. Recovery options include the following:  
  • Restore and recover:  
    - Recreate DATA disk group  
    - Restore from RECO and tape-based backups, if required  
    - Recover database  
  • Switch and recover:  
    - Use RMAN switch to copy  
    Recover database |
| Normal Redundancy for ALL (DATA and RECO) | The preceding storage outage scenarios resulted in failure of all Oracle ASM disk groups. However, using cross-disk group mirror isolation the outage is limited to one disk group. **Note:** This option is not available for eighth or quarter racks. | Oracle recommends a minimum of high redundancy for DATA only. Use the Normal Redundancy for ALL option when the primary database is protected by an Oracle Data Guard standby database deployed on a separate Oracle Exadata Database Machine or when the Exadata Database Machine is servicing only development or test databases. Oracle Data Guard provides real-time data protection and fast failover for storage failures. If Oracle Data Guard is not available and the DATA or RECO disk groups are lost, then leverage recovery options described in My Oracle Support note 1339373.1. |

The optimal file placement for setup for MAA is:

- **Oracle Database files** — DATA disk group
- **Flashback log files, archived redo files, and backup files** — RECO disk group
### Redo log files
First high redundancy disk group. If no high redundancy disk group exists, then redo log files are multiplexed across the DATA and RECO disk groups.

### Control files
First high redundancy disk group. If no high redundancy disk groups exist, the use one control file in the DATA disk group. The backup control files should reside in the RECO disk group, and `RMAN CONFIGURE CONTROLFILE AUTOBACKUP ON` should be set.

### Server parameter files (SPFILE)
First high redundancy disk group. If no high redundancy disk group exists, then SPFILE should reside in the DATA disk group. SPFILE backups should reside in the RECO disk group.

### Oracle Cluster Registry (OCR) and voting disks for Oracle Exadata Database Machine Full Rack and Oracle Exadata Database Machine Half Rack
First high redundancy disk group. If no high redundancy disk group exists, then the files should reside in the DATA disk group.

### Voting disks for Oracle Exadata Database Machine Quarter Rack or Eighth Rack
First high redundancy disk group, otherwise in normal redundancy disk group. If there are fewer than 5 Exadata storage cells with high redundancy disk group, additional quorum disks will be stored on Exadata database servers during OEDA deployment. Refer to [Oracle Exadata Database Machine Maintenance Guide](#) to migrate voting disks to high redundancy disk group for existing high redundancy disk group configurations.

### Temporary files
First normal redundancy disk group. If the high redundancy for ALL option is used, then the use the first high redundancy disk group.

### Staging and non-database files
DBFS disk group or ACFS volume

### Oracle software (including audit and diagnostic destinations)
Exadata database server local file system locations configured during OEDA deployment

#### 1.3.3 About Grid RAID
A grid Redundant Array of Independent Disks (RAID) configuration uses Oracle ASM mirroring capabilities.

To use grid RAID, you place grid disks in an Oracle ASM disk group with a normal or high redundancy level, and set all grid disks in the same cell to be in the same Oracle ASM failure group. This ensures that Oracle ASM does not mirror data extents using disks within the cell. Using disks from different cells ensures that an individual cell failure does not cause the data to be unavailable.

Grid RAID also provides simplified creation of cell disks. With grid RAID, LUNs are automatically created from available physical disks because Oracle software automatically creates the required LUNs.
1.3.4 About Storage Server Security

Security for Exadata Storage Servers is enforced by identifying which clients can access storage servers and grid disks.

Clients include Oracle ASM instances, database instances, and clusters. When creating or modifying grid disks, you can configure the Oracle ASM owner and the database clients that are allowed to use those grid disks.

Related Topics

• About Exadata Storage Server Security Modes

1.3.5 About iDB Protocol

The iDB protocol is a unique Oracle data transfer protocol that serves as the communications protocol among Oracle ASM, database instances, and storage cells.

General-purpose data transfer protocols operate only on the low-level blocks of a disk. In contrast, the iDB protocol is aware of the Oracle internal data representation and is the necessary complement to Exadata storage server specific features, such as predicate processing offload.

In addition, the iDB protocol provides interconnection bandwidth aggregation and failover.

1.3.6 About Oracle Exadata System Software Processes

Oracle Exadata System Software includes the following software processes:

• Cell Server (CELLSRV) is the primary Exadata storage server software component and provides the majority of Exadata storage services. It serves simple block requests, such as database buffer cache reads, and facilitates Smart Scan requests, such as table scans with projections and filters. CELLSRV implements Exadata I/O Resource Management (IORM), which works in conjunction with Oracle Database resource management to meter out I/O bandwidth to the various databases and consumer groups that are issuing I/Os. CELLSRV also collects numerous statistics relating to its operations. CELLSRV is a multithreaded server and typically uses the largest portion of CPU resources on a storage server.

• Offload server (CELLOFLSRV <version>) is a helper process to CELLSRV that processes offload requests from a specific database version. The offload servers enable each storage server to support all offload operations from multiple database versions. The offload servers automatically run conjunction with CELLSRV, and they require no additional configuration or maintenance.

• Management Server (MS) provides standalone Oracle Exadata System Software management and configuration functions. MS is also responsible for sending alerts, and collects some statistics in addition to those collected by CELLSRV.

• Restart Server (RS) monitors the CELLSRV, offload server, and MS processes and restarts them, if necessary.
1.3.7 About Cell Management

Each cell in the Oracle Exadata Storage Server grid is individually managed with Cell Control Command-Line Interface (CellCLI).

The CellCLI utility provides a command-line interface to the cell management functions, such as cell initial configuration, cell disk and grid disk creation, and performance monitoring. The CellCLI utility runs on the cell, and is accessible from a client computer that has network access to the storage cell or is directly connected to the cell. The CellCLI utility communicates with Management Server to administer the storage cell.

To access the cell, you should either use Secure Shell (SSH) access, or local access, for example, through a KVM switch (keyboard, video or visual display unit, mouse) switch. SSH allows remote access, but local access might be necessary during the initial configuration when the cell is not yet configured for the network. With local access, you have access to the cell operating system shell prompt and use various tools, such as the CellCLI utility, to administer the cell.

You can run the same CellCLI commands remotely on multiple cells with the dcli utility.

To manage a cell remotely from a compute node, you can use the ExaCLI utility. ExaCLI enables you to run most CellCLI commands on a cell. This is necessary if you do not have direct access to a cell to run CellCLI, or if SSH service on the cell has been disabled. To run commands on multiple cells remotely, you can use the exadcli utility.

Related Topics

- **Using the CellCLI Utility**
  You use the Cell Control Command-Line Interface (CellCLI) utility to manage Oracle Exadata System Software.

- **Using the dcli Utility**
  The dcli utility facilitates centralized management across Oracle Exadata Database Machine.

---

1.3.8 About Database Server Software

Oracle software is installed on the Exadata database servers. Oracle Exadata System Software works seamlessly with Oracle Database. The software on the database servers includes:
• Oracle Database instance, which contains the set of Oracle Database background processes that operate on the stored data and the shared allocated memory that those processes use to do their work. The database server software also includes utilities for administration, performance management, and support of the database.

• Oracle Automatic Storage Management (Oracle ASM), is a clustered file system and volume manager which provides storage management optimized for the database and Oracle Exadata Storage Servers. Oracle ASM is part of Oracle Grid Infrastructure. The Oracle Grid Infrastructure software provides the essential functions to maintain cluster coherence for all the Exadata servers. The Oracle Grid Infrastructure software also monitors the health and liveness of both database and storage servers, providing database high availability in case of planned and unplanned storage outages.

The Oracle ASM instance handles placement of data files on disks, operating as a metadata manager. The Oracle ASM instance is primarily active during file creation and extension, or during disk rebalancing following a configuration change. Run-time I/O operations are sent directly from the database to storage cells without passing through an Oracle ASM instance.

• The Oracle Database Resource Manager, which ensures that I/O resources are properly allocated within a database.

• The iDB protocol is used by the database instance to communicate with cells, and is implemented in an Oracle-supplied library statically linked with the database server.

Related Topics

• Managing I/O Resources
  I/O Resource Management (IORM) is a tool for managing how multiple databases and the workloads within the databases share the I/O resources of Oracle Exadata System Software.

• About iDB Protocol
  The iDB protocol is a unique Oracle data transfer protocol that serves as the communications protocol among Oracle ASM, database instances, and storage cells.

1.3.9 About Oracle Enterprise Manager for Oracle Exadata Database Machine

Oracle Enterprise Manager provides a complete target that enables you to monitor Oracle Exadata Database Machine, including configuration and performance, in a graphical user interface (GUI).

The following figure shows the Exadata Storage Server Grid home page. Viewing this page, you can quickly see the health of the storage servers, key storage performance characteristics, and resource utilization of storage by individual databases.
In addition to reports, Oracle Enterprise Manager enables you to set metric thresholds for alerts and monitor metric values to determine the health of your Exadata systems.

Related Topics

- Oracle Enterprise Manager Exadata Management Getting Started Guide
This chapter describes the major steps to configure Oracle Exadata System Software.

You determine the number of disks and cells needed in the grid based on your requirements for capacity, performance, and redundancy.

Hardware and software have already been installed for the storage servers. The procedures in this chapter describe how to configure a storage cell for use with the Oracle Database and Oracle Automatic Storage Management (Oracle ASM) instances.

Note:
Modifications to the Oracle Exadata Storage Server hardware or software are not supported. Only the documented network interfaces on the Oracle Exadata Storage Server should be used for all connectivity including management and storage traffic. Additional network interfaces should not be used.

- **Understanding Oracle Exadata System Software Release Numbering**
  The Oracle Exadata System Software release number is related to the Oracle Database release number.

- **Understanding Oracle Exadata Storage Server Configuration**
  Oracle Exadata Storage Server ships with all hardware and software pre-installed, however you must configure Oracle Exadata System Software for your environment.

- **Network Configuration and IP Addresses Recommendations**
  Follow the recommendation network configuration for Oracle Exadata Storage Server.

- **Assigning IP Addresses for Oracle Exadata Database Machine**
  This topic summarizes the Oracle Exadata Database Machine network preparation before installing the new storage server.

- **Configuring Oracle Exadata System Software for Your Location**
  Customize the software installation for your environment.

- **Configuring Cells, Cell Disks, and Grid Disks with CellCLI**
  Configure the cells, cell disks and grid disks for each new Oracle Exadata Storage Server.

- **Creating Flash Cache and Flash Grid Disks**
  Oracle Exadata Storage Servers are equipped with flash disks. These flash disks can be used to create flash grid disks to store frequently accessed data.
2.1 Understanding Oracle Exadata System Software Release Numbering

The Oracle Exadata System Software release number is related to the Oracle Database release number.

The Oracle Exadata System Software release number matches the highest Oracle Grid Infrastructure and Oracle Database version it supports. For example, the highest version Oracle Exadata System Software release 18 supports is Oracle Grid Infrastructure and Oracle Database release 18. The highest version Oracle Exadata System Software release 12.2 supports is Oracle Grid Infrastructure and Oracle Database release 12.2.0.1.

Release 18c and Later Numbering

The Oracle Exadata System Software release that followed release 12.2.1.1.8 was renamed to 18.1.0 and a new numbering scheme for the Oracle Exadata System Software was implemented. Instead of a legacy nomenclature such as 12.2.1.1.5, a three field format consisting of: Year.Update.Revision is used, for example 18.1.0. This new numbering scheme allows you to clearly determine:

- The annual release designation of the software
- The latest software update, which can contain new features
- The latest software revision, which includes security and software fixes

If there are new features or new hardware supported, a new software update will be release during the year, for example, 19.2. To allow you to keep current on just security-related and other software fixes after your feature environment becomes stable, software revisions are made available approximately once a month, for example 19.1.3.

Numbering for Releases Prior to 18c

- The first two digits of the Oracle Exadata System Software release number represent the major Oracle Database release number, such as Oracle Database 12c Release 1 (12.1). Oracle Exadata System Software release 12.1 is compatible with all Oracle Database 12c Release 1 (12.1) releases.
- The third digit usually represents the component-specific Oracle Database release number. This digit usually matches the fourth digit of the complete release number, such as 12.1.0.1.0 for the current release of Oracle Database.
- The last two digits represent the Oracle Exadata System Software release.
2.2 Understanding Oracle Exadata Storage Server Configuration

Oracle Exadata Storage Server ships with all hardware and software pre-installed, however you must configure Oracle Exadata System Software for your environment.

This topic provides a general overview of the configuration tasks. Subsequent topics describe the actual procedures.

- **Assign IP Addresses for the Storage Servers**
  As part of configuring the storage servers, you assign IP addresses to connect the storage to the various networks.

- **Configure the Storage Server for Your Location**
  Configure the storage server for use within your company.

- **Configure the Storage Cell**
  Use the `ALTER CELL` command to configure the cell.

- **Verify Storage Cell Attributes**
  Use the `LIST CELL DETAIL` command to verify the storage cell attributes.

- **Create the Storage Cell Disks**
  Use the `CREATE CELLDISK` command to create the cell disks.

- **Create the Grid Disks**
  Use the `CREATE GRIDDISK` command to create the grid disks. The size of the disks depends on your requirements.

- **Create the PMEM Cache**
  By default, the `CREATE CELL` command creates the cell disks, which are used when creating the PMEM Cache.

- **Create the Flash Disks and Flash Cache**
  By default, the `CREATE CELL` command creates flash cell disks on all flash disks. The command then creates Exadata Smart Flash Cache on the flash cell disks.

- **Configure Oracle Auto Service Request (ASR)**
  Oracle Auto Service Request (ASR) for Oracle Exadata Database Machine automatically creates service requests by detecting common hardware faults.

2.2.1 Assign IP Addresses for the Storage Servers

As part of configuring the storage servers, you assign IP addresses to connect the storage to the various networks.

Assign IP addresses for each storage server for the following ports:

- Network access port
- Remote management port
- RDMA Network Fabric port
2.2.2 Configure the Storage Server for Your Location

Configure the storage server for use within your company.

1. Power on the storage server.
2. Change the default passwords.
3. Set the time zone on the storage server to make the local time.
4. Configure other information as needed, such as NTP and DNS servers.

Related Topics

- Understanding User Accounts
- Change the DNS Server on Exadata Storage Server
- Change Time Zone Settings on Exadata Storage Servers
- Set the NTP Server on Exadata Storage Servers

2.2.3 Configure the Storage Cell

Use the `ALTER CELL` command to configure the cell.

In Example 2-1, e-mail notification is configured to send notification messages to the storage server administrator according to the specified notification policy. The hyphen (-) at the end of each line of the `ALTER CELL` command allows the command to continue to additional lines before pressing Enter. As an alternative, you can run the command using a text file.

**Example 2-1 Configuring a New Cell**

```sql
CellCLI> ALTER CELL                                      -
smtpServer='my_mail.example.com',               -
smtpFromAddr='john.doe@example.com',            -
smtpToAddr='jane.smith@example.com',           -
notificationPolicy='clear',                    -
notificationMethod='mail,snmp'
```

2.2.4 Verify Storage Cell Attributes

Use the `LIST CELL DETAIL` command to verify the storage cell attributes.

**Example 2-2 Viewing Storage Cell Details**

This example shows how to view the storage cell attributes.

```sql
CellCLI> LIST CELL DETAIL
name:                   cell01
accessLevelPerm:        remoteLoginEnabled
bbuStatus:              normal
cellVersion:            OSS_19.3.0.0.0_LINUX.X64_190910
cpuCount:               64/64
diagHistoryDays:        7
doNotServiceLEDStatus:  off
```
fanCount: 12/12  
fanStatus: normal  
flashCacheMode: WriteBack  
httpsAccess: ALL  
id: 1904XCA016  
interconnectCount: 2  
interconnect1: ib0  
interconnect2: ib1  
iormBoost: 0.0  
ipaddress1: 192.168.41.245/21  
ipaddress2: 192.168.41.246/21  
kernelVersion: 4.14.35-1902.5.0.el7uek.x86_64  
locatorLEDStatus: off  
makeModel: Oracle Corporation ORACLE SERVER X8-2L  
High Capacity  
memoryGB: 188  
metricHistoryDays: 7  
offloadGroupEvents:  
pmemCacheMode: WriteThrough  
powerCount: 2/2  
powerStatus: normal  
ramCacheMaxSize: 0  
ramCacheMode: On  
ramCacheSize: 0  
releaseImageStatus: success  
releaseVersion: 19.3.0.0.0.190824  
rpmVersion: cell-19.3.0.0.0_LINUX.X64_190824-1.x86_64  
releaseTrackingBug: 29344484  
smtpFrom: "Exadata Admins"  
smtpFromAddr: exa_admins@example.com  
smtpServer: my_mail_svr.example.com  
smtpToAddr: jane.smith@example.com  
snmpSubscriber: host=host1,port=162,community=public,type=asr,asrmPort=16161  
status: online  
temperatureReading: 29.0  
temperatureStatus: normal  
upTime: 2 days, 7:05  
usbStatus: normal  
cellsrvStatus: stopped  
msStatus: running  
rsStatus: running

2.2.5 Create the Storage Cell Disks

Use the CREATE CELLDISK command to create the cell disks.

In Example 2-3, the ALL option creates all the cell disks using the default names.

The cell disks are created with names in the form CD_lunID_cellname. The lunID and cellname values correspond to the id attribute of the LUN and name attribute of the cell. You can specify other disk names if you create single cell disks.
On Oracle Exadata Storage Server with flash disks, the `CREATE CELLDISK ALL` command also creates cell disks on the flash disks.

```
CellCLI> CREATE CELLDISK ALL
CellDisk FD_01_cell01 successfully created
CellDisk FD_02_cell01 successfully created
CellDisk FD_03_cell01 successfully created
CellDisk FD_04_cell01 successfully created
CellDisk FD_05_cell01 successfully created
CellDisk FD_06_cell01 successfully created
CellDisk FD_07_cell01 successfully created
CellDisk FD_08_cell01 successfully created
CellDisk FD_09_cell01 successfully created
CellDisk FD_10_cell01 successfully created
CellDisk FD_11_cell01 successfully created
CellDisk FD_12_cell01 successfully created
CellDisk FD_13_cell01 successfully created
CellDisk FD_14_cell01 successfully created
CellDisk FD_15_cell01 successfully created
```

**Note:**
The `CREATE CELLDISK` command creates cell disks on flash disks if they do not currently exist. If there are cell disks on the flash disks, then they are not created again.

**Example 2-3  Creating Cell Disks**

```
CellCLI> CREATE CELLDISK ALL
CellDisk CD_00_cell01 successfully created
CellDisk CD_01_cell01 successfully created
CellDisk CD_02_cell01 successfully created
...
CellDisk CD_10_cell01 successfully created
CellDisk CD_11_cell01 successfully created
```

### 2.2.6 Create the Grid Disks

Use the `CREATE GRIDDISK` command to create the grid disks. The size of the disks depends on your requirements.

1. Determine the naming format for the grid disks or use the `ALL PREFIX` option.

Grid disk names must be unique across all cells within a single deployment. By following the recommended naming conventions for naming the grid and cell disks you automatically get unique names. If you do not use the default generated name when creating grid disks, then you must ensure that the grid disk name is unique across all storage cells. If the disk name is not unique, then it might not be possible to add the grid disk to an Oracle Automatic Storage Management (Oracle ASM) disk group.
When the ALL PREFIX option is used, the generated grid disk names are composed of the grid disk prefix followed by an underscore (_) and then the cell disk name.

2. Use the CREATE GRIDDISK command to create the grid disks.

When creating a grid disk:

- You do not have to specify the size attribute. The maximum size possible is automatically chosen if the size attribute is omitted.

- For rotating disks, offset determines the position on the disk where the grid disk is allocated. The outermost tracks have lower offset values, and these tracks have greater speed and higher bandwidth. Offset can be explicitly specified to create grid disks that are relatively higher performing than other grid disks. If offset is not specified, then the best (warmest) available offset is chosen automatically in chronological order of grid disk creation. You should first create those grid disks expected to contain the most frequently accessed (hottest) data, and then create the grid disks that contain the relatively colder data.

Note:
Offset does not apply to Flash Disk.

- Sparse grid disks only need to be created when using snapshots. The sparse disk stores the files generated by the snapshot. All standard grid disk operations are supported for sparse grid disks. Sparse grid disks have an additional attribute, virtualsize. The attribute configures the maximum virtual space the grid disk must provide. The attribute can be resized if the configuration runs out of virtual space on the sparse grid disk and there is physical space available.

The maximum allowed size of a sparse disk is the size of free space on the cell disk. The maximum allowed virtual size is 100 TB.

Oracle Exadata System Software monitors physical space used by sparse grid disks, and generates an alert when most of the space is used. To avoid out-of-space errors, add more physical space to the grid disk using the ALTER GRIDDISK command, or delete some of the Oracle ASM files to free space on the grid disk.

Example 2-4 Creating Grid Disks

This example shows how to create grid disks. In this example, the ALL HARDDISK PREFIX option creates one grid disk on each cell disk of the storage cell. The Oracle ASM disk group name is used with PREFIX to identify which grid disk belongs to the disk group. Prefix values data and reco are the names of the Oracle ASM disk groups that are created.

CellCLI> CREATE GRIDDISK ALL HARDDISK PREFIX=data, size=300G
GridDisk data_CD_00_cell01 successfully created
GridDisk data_CD_01_cell01 successfully created
GridDisk data_CD_02_cell01 successfully created
...
GridDisk data_CD_11_cell01 successfully created
CellCLI> CREATE GRIDDISK ALL HARDDISK PREFIX=reco, size=600G
GridDisk reco_CD_00_cell01 successfully created
GridDisk reco_CD_01_cell01 successfully created
GridDisk reco_CD_02_cell01 successfully created
... GridDisk reco_CD_11_cell01 successfully created

The LIST GRIDDISK command shows the grid disks that are created.

CellCLI> LIST GRIDDISK
  data_CD_00_cell01       active
  data_CD_01_cell01       active
  data_CD_02_cell01       active
  ...
  data_CD_11_cell01       active
  reco_CD_00_cell01      active
  reco_CD_01_cell01      active
  reco_CD_02_cell01      active
  ...
  reco_CD_11_cell01      active

Example 2-5    Creating a Sparse Grid Disk

In this example, the sparse grid disk uses up to 300 GB from the physical cell disk size, but it exposes 20000 GB virtual space for the Oracle ASM files.

CellCLI> CREATE GRIDDISK ALL HARDDISK PREFIX=sp, size=300G, virtualsize=20000G
GridDisk sp_CD_00_cell01 successfully created
GridDisk sp_CD_01_cell01 successfully created
GridDisk sp_CD_02_cell01 successfully created
... GridDisk sp_CD_11_cell01 successfully created

2.2.7 Create the PMEM Cache

By default, the CREATE CELL command creates the cell disks, which are used when creating the PMEM Cache.

- Use the CREATE PMEMCACHE ALL to create the PMEM Cache on all PMEM cell disks.

You can use the ALTER PMEMCACHE command to alter the set of cell disks used by PMEM cache, flush dirty blocks from PMEM cache, or cancel a previous flush operation on the specified cell disks to re-enable caching.

2.2.8 Create the Flash Disks and Flash Cache

By default, the CREATE CELL command creates flash cell disks on all flash disks. The command then creates Exadata Smart Flash Cache on the flash cell disks.

- Use the CREATE GRIDDISK ALL FLASHDISK PREFIX='FLASH' and CREATE FLASHCACHE commands to create the flash disks and flash cache.
To change the size of the Exadata Smart Flash Cache or create flash grid disks it is necessary to remove the flash cache, and then create the flash cache with a different size, or create the flash grid disks.

### 2.2.9 Configure Oracle Auto Service Request (ASR)

Oracle Auto Service Request (ASR) for Oracle Exadata Database Machine automatically creates service requests by detecting common hardware faults.

ASR support covers selected components, such as disks and flash cards, in Oracle Exadata Storage Servers and Oracle Exadata Database Servers.

- If you did not elect to configure Oracle Auto Service Request (ASR) when using Oracle Exadata Deployment Assistant (OEDA) to configure your Oracle Exadata Rack, then refer to [Oracle Auto Service Request Quick Installation Guide for Oracle Exadata Database Machine](#) for configuration instructions.

#### Related Topics
- [Oracle Exadata Database Machine Maintenance Guide](#)
- About Oracle Auto Service Request

### 2.3 Network Configuration and IP Addresses Recommendations

Follow the recommendation network configuration for Oracle Exadata Storage Server.

- If your network is not already configured, then set up a fault-tolerant, private network subnet for Oracle Exadata Database Machine storage servers and database servers with multiple switches to eliminate the switch as a single point of failure. If all the interconnections in the storage network are connected through a single switch, then that switch can be a single point of failure.

If you are using a managed switch, then ensure that the switch VLAN configuration isolates storage server network traffic from all other network traffic.

- Allocate a block of IP addresses for the storage server general administration and the Integrated Lights Out Manager (ILOM) interfaces. Typically, these interfaces are on the same subnet, and may share the subnet with other hosts. For example, on the 192.168.200.0/24 subnet, you could assign the block of IP addresses between 192.168.200.31 and 192.168.200.100 for storage server general administration and ILOM interfaces. Other hosts sharing the subnet would be allocated IP addresses outside the block. The general administration and ILOM interfaces can be placed on separate subnets, but this is not required.

The following is a sample of four non-overlapping blocks of addresses. One set of addresses should be assigned to the normal Gigabit Ethernet interface for the storage servers. The other blocks can be assigned for the ILOM port for the storage servers. The third set can be used for the database server Gigabit Ethernet ports, and the fourth for the database server ILOM ports.

```
192.168.200.0/21 (netmask 255.255.248.0)
192.168.208.0/21 (netmask 255.255.248.0)
192.168.216.0/21 (netmask 255.255.248.0)
192.168.224.0/21 (netmask 255.255.248.0)
```
• The RDMA Network Fabric network should be a private network for use by the database server hosts and storage servers, and can have private local network addresses. These addresses must also be allocated in non-overlapping blocks.

The following example has 2 blocks of local RDMA Network Fabric addresses. Both the database server RDMA Network Fabric addresses and the storage server RDMA Network Fabric addresses must be in the same subnet to communicate with each other. With bonding, only one subnet is necessary for RDMA Network Fabric addresses.

192.168.50.0/24 (netmask 255.255.255.0)
192.168.51.0/24 (netmask 255.255.255.0)

The preceding subnet blocks do not conflict with each other, and do not conflict with the current allocation to any of the hosts. It is a good practice to allocate the subnet blocks so that they have an identical netmask, which helps to simplify network administration.

• For Oracle Exadata Database Machine, the maximum allowed number of hosts in an RDMA Network Fabric network is 4096. Therefore, the network prefix value for the RDMA Network Fabric network must be equal to or greater than 20. This means the netmask must be between 255.255.240.0 and 255.255.255.254 inclusive.

You can determine the network prefix value for a given host IP address and its netmask using the `ipcalc` utility on any Oracle Linux machine, as follows:

```
ipcalc <host ip address> -m <netmask for the host> -p
```

• Do not allocate addresses that end in .0, .1, or .255, or those that would be used as broadcast addresses for the specific netmask that you have selected. For example, avoid addresses such as 192.168.200.0, 192.168.200.1, and 192.168.200.255.

• Ensure the network allows for future expansion. For example, 255.255.255.254 is valid network (prefix /31) but it only allows 1 host.

• If a domain name system (DNS) is required, then set up your DNS to help reference storage servers and interconnections. Oracle Exadata Database Machine storage servers do not require DNS. However, if DNS is required, then set up your DNS with the appropriate IP address and host name of the storage servers.

• The RDMA Network Fabric network should be used for network and storage communication when using Oracle Clusterware. Use the following command to verify the private network for Oracle Clusterware communication is using the RDMA Network Fabric:

```
oifcfg getif -type cluster_interconnect
```

• The Reliable Data Socket (RDS) protocol should be used over the RDMA Network Fabric network for database server to storage server communication and Oracle Real Application Clusters (Oracle RAC) communication. Check the alert log to verify the private network for Oracle RAC is running the RDS protocol over the
RDMA Network Fabric network. The following message should be in the alert logs for the instances:

```
class interconnect IPC version: Oracle RDS/IP (generic)
```

If the RDS protocol is not being used over the RDMA Network Fabric network, then perform the following procedure:

1. Shut down any processes that are using the Oracle software.
2. Change to the `ORACLE_HOME/rdbms/lib` directory.
3. Relink the Oracle Database software with the RDS protocol.

   ```
   make -f ins_rdbms.mk ipc_rds ioracle
   ```

   **Note:**
   
   If Oracle ASM uses a separate Oracle home from the database instance, then RDS should be enabled for the binaries in both homes.

---

### 2.4 Assigning IP Addresses for Oracle Exadata Database Machine

This topic summarizes the Oracle Exadata Database Machine network preparation before installing the new storage server.

Each storage server contains the following network ports:

- One dual-port RDMA Network Fabric card
  
  Oracle Exadata Storage Servers are designed to be connected to two separate RDMA Network Fabric switches for high availability. The dual port card is only for availability. Each port of the RDMA Network Fabric card is capable of transferring the full data bandwidth generated by the storage server. The loss of one network connection does not impact the performance of the storage server.

- Ethernet ports for normal network access, depending on the platform
  
  - Oracle Exadata Database Machine X6-2 and earlier storage servers come with four Ethernet ports. However, only connect one port to a switch, and configure it for network access.

- One Ethernet port is exposed by the Baseboard Management Controller (BMC), or Management Controller (MC) on the storage server. This port is used for Integrated Lights Out Manager (ILOM).

To prepare the storage server for network access, you must perform the following steps:

1. Assign one address to the bonded RDMA Network Fabric port. When you first set up the storage server, you are prompted for the RDMA Network Fabric configuration information. This information is used automatically during the `CREATE CELL` command on first boot, and provides the data path for communication between the storage server and the database servers.
To change the IP address after the initial configuration, use the following command, where *interface_name* is the interface name for the RDMA Network Fabric:

```
CREATE CELL interconnect1=interface_name,
interconnect2=interface_name
```

For InfiniBand Network Fabric networks, the interface names are `ib0` and `ib1`. For RoCE Network Fabric networks, the interface names are `re0` and `re1`.

2. Assign an IP address to the storage server for network access.

3. Assign an IP address to the storage server for ILOM.

You can access the remote management functionality with a Java-enabled Web browser at the assigned IP address.

---

**See Also:**


---

### 2.5 Configuring Oracle Exadata System Software for Your Location

Customize the software installation for your environment.

- **Configuring ILOM With Static IP for Oracle Exadata Storage Servers**
  
  Basic lights-out remote management configuration is done during the first boot.

- **Preparing the Servers**

  Use the following steps to prepare the database servers and storage servers for use.

---

**Caution:**

Do not enable the sideband management available in ILOM. Doing so disables all the SNMP agent reporting and monitoring functionality for the server.
2.5.2 Preparing the Servers

Use the following steps to prepare the database servers and storage servers for use.

2. Power on the storage server to boot its operating system.
3. Respond to the prompts to configure the system, after the storage server boots.
   Press \texttt{y} to confirm, or \texttt{n} to retry or terminate when you are prompted for a yes or no response during the configuration steps. The yes or no prompt shows the default choice in brackets. If you do not enter a response, then the default choice is selected when you press \texttt{Enter}.
4. Check the network connections.
   The list of all discovered interfaces displays, and you are prompted to check the cables for those interfaces that do not have an active network cable connection. You can retry the configuration steps after connecting the cables, or ignore the unconnected interfaces. Only connected interfaces can be configured.
5. Enter the DNS server IP addresses, if needed.
   A DNS is not needed for a standalone, private storage environment.
6. Enter the time preference.
   • Choose the local time region number from the displayed list of available time regions.
   • Choose the location within the time region number from the displayed list of locations.
7. Enter the Network Time Protocol (NTP) servers.
   These servers are required to maintain the time on the system correctly, and are synchronized to your local time source.
8. Enter the Ethernet addresses, RDMA Network Fabric IP addresses and interfaces.
   A list of all Ethernet and RDMA Network Fabric interfaces that have an active network connection is displayed with the name of the interface on the extreme left. The InfiniBand Network Fabric interface is named \texttt{BONDIB0} and uses bonding between physical InfiniBand Network Fabric interfaces \texttt{ib0} and \texttt{ib1}. Bonding provides the ability to transparently fail over from \texttt{ib0} to \texttt{ib1} or from \texttt{ib1} to \texttt{ib0} if connectivity is lost to \texttt{ib0} or \texttt{ib1}, respectively.
   For each Ethernet and RDMA Network Fabric interface you configure, you are prompted for the following that apply to the interface:
   • IP address
   • Netmask
   • Gateway IP address
   • Fully-qualified domain name
   If you choose not to configure each interface in the list, then that interface is not configured, and it does not start at system startup. After the configuration of the IP addresses, the system completes the startup process. At the end of the process, additional packages are installed.
9. Select the canonical, fully-qualified domain name from the list.

This host name is the primary public host name for the server, and is part of the /etc/sysconfig/network file.

If more than one Ethernet interface was configured with the gateway, then select the line number for the default gateway. This gateway is in the /etc/sysconfig/network file, and is used as the default gateway.

10. Provide the following information when prompted for it:

- ILOM full, domain-qualified host name
- ILOM IP address
- ILOM netmask
- ILOM gateway
- ILOM NTP servers
- ILOM DNS server

11. (Oracle Exadata Storage Server only) Change the initial passwords for the root, celladmin, and cellmonitor users to more secure passwords.

    ![Note:]

    If you do not have the password for the root user, then contact Oracle Support Services.

To change the passwords, log in as the root user, then use the passwd command to change the passwords, such as the following:

```bash
# passwd
# passwd celladmin
# passwd cellmonitor
```

To verify the changed passwords, log in and out of the server using each of the user names.

![Note:]

The cellmonitor user is set up with privileges that enable you to only view Oracle Exadata Storage Server objects. You must be logged in as the celladmin user to perform administrative tasks.

12. Check for any failures reported in the /var/log/cellos/vldrun.first_boot.log file after the first boot configuration.

For each failed validation, perform the following procedure:


    The file exists only if the validation process has identified some corrective action. Follow the suggestions in the file to correct the cause of the failure.
b. If the SuggestedRemedy file does not exist, then examine the log file for the failed validation in /var/log/cellos/validations to track down the cause, and correct it as needed.

13. (Oracle Exadata Storage Server only) Use the following commands to verify acceptable performance levels:

```
# cellcli -e "alter cell shutdown services cellsrv"
# cellcli -e "calibrate"
```

Related Topics

- **Configuring iLOM With Static IP for Oracle Exadata Storage Servers**
  
  Basic lights-out remote management configuration is done during the first boot.

- **Using the ipconf Utility**
  
  The ipconf utility is used to set and change the following parameters on Oracle Exadata servers.

---

**2.6 Configuring Cells, Cell Disks, and Grid Disks with CellCLI**

Configure the cells, cell disks and grid disks for each new Oracle Exadata Storage Server.

The steps described in Preparing the Servers must be completed before starting this task.

During the procedure, you can display help using the HELP command, and object attributes using the DESCRIBE command. Example 2-6 shows how to display help and a list of attributes for Oracle Exadata Database Machine CELL objects.

1. Log in as the celladmin user.

2. Use the cellcli command to start Cell Control Command-Line Interface (CellCLI) to connect to the storage cell.

   The required cell services, Restart Server (RS) and Management Server (MS), should be running after the binary has been installed. If not, then an error message displays when using the CellCLI utility. If an error message displays, then run the following commands to start Oracle Exadata System Software RS and MS services:

   ```
   CellCLI> ALTER CELL STARTUP SERVICES RS
   CellCLI> ALTER CELL STARTUP SERVICES MS
   ```

3. Configure the cell using the CellCLI ALTER CELL command. During first boot, the cell is created, and the flash cell disks and flash cache defined automatically.

   The following example shows how to use the ALTER CELL command to set the cell name and configure e-mail notification, which sends notification messages to the storage server administrator according to the specified notification policy.

   ```
   CellCLI> ALTER CELL
   name=cell_name,                                      -
   ```
smtpServer='my_mail.example.com',
smtplFromAddr='john.doe@example.com',
smtplToAddr='jane.smith@example.com',
notificationPolicy='clear',
notificationMethod='mail,snmp'

4. Use the following command to check the storage cell attributes, and to verify the current configuration:

   CellCLI> LIST CELL DETAIL

5. Create the cell disks, using the CREATE CELLDISK command. In most cases, you can use the default cell disk names and LUN IDs. Use the following command to create cell disks and LUN IDs with the default values.

   CellCLI> CREATE CELLDISK ALL

6. Create grid disks on each cell disk of the storage cell, using the CREATE GRIDDISK command.

7. Exit the CellCLI utility after setting up the storage cell using the following command:

   CellCLI> EXIT

8. Repeat the configuration process for each new storage cell. This procedure must be done on each new cell before configuring the Exadata Cell realm, the database server hosts, or the database and Oracle ASM instances.

Example 2-6  Displaying Help Information

   CellCLI> HELP
   CellCLI> HELP CREATE CELL
   CellCLI> HELP ALTER CELL
   CellCLI> DESCRIBE CELL

After you complete the cell configuration, you can perform the following optional steps on the storage cell:

• Add the storage cell to the Exadata Cell realm.

• Configure security on the Oracle Exadata Storage Server grid disks, as described in Configuring Security for Oracle Exadata Storage Server Software.

• Configure an inter-database plan for a cell rather than using the default plans, as described in Managing I/O Resources.

For database server hosts other than those in Oracle Exadata Database Machine, refer to release notes for enabling them to work with Oracle Exadata Storage Servers.

Related Topics

• ALTER CELL
Oracle Exadata Storage Servers are equipped with flash disks. These flash disks can be used to create flash grid disks to store frequently accessed data.

Alternatively, all or part of the flash disk space can be dedicated to Exadata Smart Flash Cache. In this case, the most frequently-accessed data is cached in Exadata Smart Flash Cache.

The `ALTER CELLDISK ... FLUSH` command must be run before exporting a cell disk to ensure that the data not synchronized with the disk (dirty data) is flushed from flash cache to the grid disks.

- By default, the `CREATE CELL` command creates flash cell disks on all flash disks, and then creates Exadata Smart Flash Cache on the flash cell disks.
  
  To change the size of the Exadata Smart Flash Cache or create flash grid disks it is necessary to remove the flash cache, and then create the flash cache with a different size, or create the flash grid disks.

- To change the amount of flash cache allocated, use the `flashcache` attribute with the `CREATE CELL` command.
  
  If the `flashcache` attribute is not specified, then all available flash space is allocated for flash cache.

- To explicitly create the Exadata Smart Flash Cache, use the `CREATE FLASHCACHE` command. Use the `celldisk` attribute to specify which flash cell disks contain the Exadata Smart Flash Cache.
  
  Alternatively, you can specify `ALL` instead of `celldisk` to use all flash cell disks. Use the `size` attribute to specify the total size of the flash cache to allocate. The allocation is evenly distributed across all flash cell disks.

**Example 2-7 Using the CREATE FLASHCACHE Command**

This example shows how to create the Exadata Smart Flash Cache. The entire size of the flash cell disk is not used because the `size` attribute has been set.

```
CellCLI> CREATE FLASHCACHE ALL size=100g
Flash cache cell101_FLASHCACHE successfully created
```

**Example 2-8 Using the CREATE GRIDDISK Command to Create Flash Grid Disks**

This example shows how to use the remaining space on the flash cell disks to create flash grid disks.

```
CellCLI> CREATE GRIDDISK ALL FLASHDISK PREFIX='FLASH'
GridDisk FLASH_FD_00_cell101 successfully created
GridDisk FLASH_FD_01_cell101 successfully created
GridDisk FLASH_FD_02_cell101 successfully created
```
GridDisk FLASH_FD_03_cell01 successfully created
GridDisk FLASH_FD_04_cell01 successfully created
GridDisk FLASH_FD_05_cell01 successfully created
GridDisk FLASH_FD_06_cell01 successfully created
GridDisk FLASH_FD_07_cell01 successfully created
GridDisk FLASH_FD_08_cell01 successfully created
GridDisk FLASH_FD_09_cell01 successfully created
GridDisk FLASH_FD_10_cell01 successfully created
GridDisk FLASH_FD_11_cell01 successfully created
GridDisk FLASH_FD_12_cell01 successfully created
GridDisk FLASH_FD_13_cell01 successfully created
GridDisk FLASH_FD_14_cell01 successfully created
GridDisk FLASH_FD_15_cell01 successfully created

CellCLI> LIST GRIDDISK
FLASH_FD_00_cell01      active
FLASH_FD_01_cell01      active
FLASH_FD_02_cell01      active
FLASH_FD_03_cell01      active
FLASH_FD_04_cell01      active
FLASH_FD_05_cell01      active
FLASH_FD_06_cell01      active
FLASH_FD_07_cell01      active
FLASH_FD_08_cell01      active
FLASH_FD_09_cell01      active
FLASH_FD_10_cell01      active
FLASH_FD_11_cell01      active
FLASH_FD_12_cell01      active
FLASH_FD_13_cell01      active
FLASH_FD_14_cell01      active
FLASH_FD_15_cell01      active

Example 2-9  Displaying the Exadata Smart Flash Cache Configuration for a Cell

Use the LIST FLASHCACHE command to display the Exadata Smart Flash Cache configuration for the cell, as shown in this example.

CellCLI> LIST FLASHCACHE DETAIL
name:                   cell01_FLASHCACHE
 cellDisk:               FD_00_cell01,FD_01_cell01,FD_02_cell01,
                         FD_03_cell01,FD_04_cell01,FD_05_cell01,
                         FD_06_cell01,FD_07_cell01,FD_08_cell01,
                         FD_09_cell01,FD_10_cell01,FD_11_cell01,
                         FD_12_cell01,FD_13_cell01,FD_14_cell01,
                         FD_15_cell01
creationTime:           2009-10-19T17:18:35-07:00
id:                     b79b3376-7b89-4de8-8051-6eefc442c2fa
Example 2-10  Dropping Exadata Smart Flash Cache from a Cell

To remove Exadata Smart Flash Cache from a cell, use the `DROP FLASHCACHE` command.

```
CellCLI> DROP FLASHCACHE
Flash cache cell01_FLASHCACHE successfully dropped
```

Related Topics

- CREATE CELL
- CREATE GRIDDISK
- LIST GRIDDISK
- LIST FLASHCACHE
- DROP FLASHCACHE

2.8 Setting Up Configuration Files for a Database Server Host

After Oracle Exadata Storage Server is configured, the database server host must be configured with the `cellinit.ora` and the `cellip.ora` files to use the cell.

The files are located in the `/etc/oracle/cell/network-config` directory.

- The `cellinit.ora` file contains the database IP addresses.
- The `cellip.ora` file contains the storage cell IP addresses.

Both files are located on the database server host. These configuration files contain IP addresses, not host names.

The `cellinit.ora` file is host-specific, and contains all database IP addresses that connect to the storage network used by Oracle Exadata Storage Servers. This file must exist for each database that connects to Oracle Exadata Storage Servers. The IP addresses are specified in Classless Inter-Domain Routing (CIDR) format. The first IP address must be designated as `ipaddress1`, the second IP address as `ipaddress2`, and so on.

The following list is an example of the IP address entry for a single database server in Oracle Exadata Database Machine:

- Oracle Exadata Database Server in Oracle Exadata Database Machine X4-2
  - `ipaddress1=192.168.10.1/22`
  - `ipaddress2=192.168.10.2/22`
- Oracle Exadata Database Server in Oracle Exadata Database Machine X3-2 or Oracle Exadata Database Machine X2-2
  - `ipaddress1=192.168.50.23/24`
• Oracle Exadata Database Server in Oracle Exadata Database Machine X3-8 or Oracle Exadata Database Machine X2-8

  – ipaddress1=192.168.41.111/21
  – ipaddress2=192.168.41.112/21
  – ipaddress3=192.168.41.113/21
  – ipaddress4=192.168.41.114/21

The IP addresses should not be changed after this file is created.

Note:

At boot time on an 8-socket system, each database server generates a cellaffinity.ora configuration file. The cellaffinity.ora file resides in the /etc/oracle/cell/network-config directory, and must be readable by Oracle Database.

The file contains a mapping between the NUMA node numbers and the IP address of the network interface card closest to each server. Oracle Database uses the file to select the closest network interface card when communicating with Oracle Exadata Storage Servers, thereby optimizing performance.

This file is only generated and used on an 8-socket system. On a 2-socket system, there is no performance to be gained in this manner, and no cellaffinity.ora file. The file is not intended to be directly edited with a text editor.

To configure a database server host for use with a cell, refer to Maintaining the RDMA Network Fabric for RoCE Network or Maintaining the RDMA Network Fabric for InfiniBand Network.

2.9 Understanding Automated Cell Maintenance

The Management Server (MS) includes a file deletion policy based on the date. When there is a shortage of space in the Automatic Diagnostic Repository (ADR) directory, then MS deletes the following files:

• All files in the ADR base directory older than 7 days.
• All files in the LOG_HOME directory older than 7 days.
• All metric history files older than 7 days.

The retention period of seven days is the default. The retention period can be modified using the metricHistoryDays and diagHistoryDays attributes with the ALTER CELL command. The diagHistoryDays attribute controls the ADR files, and the metricHistoryDays attribute controls the other files.

If there is sufficient disk space, then trace files are not purged. This can result in files persisting in the ADR base directory past the time limit specified by diagHistoryDays.
In addition, the alert.log file is renamed if it is larger than 10 MB, and versions of the file that are older than 7 days are deleted if their total size is greater than 50 MB.

MS includes a file deletion policy that is triggered when file system utilization is high. Deletion of files in the / (root) directory and the /var/log/oracle directory is triggered when file utilization is 80 percent. Deletion of files in the /opt/oracle file system is triggered when file utilization reaches 90 percent, and the alert is cleared when utilization is below 85 percent. An alert is sent before the deletion begins. The alert includes the name of the directory, and space usage for the subdirectories. In particular, the deletion policy is as follows:

- The /var/log/oracle file systems, files in the ADR base directory, metric history directory, and LOG_HOME directory are deleted using a policy based on the file modification time stamp.
  - Files older than the number of days set by the metricHistoryDays attribute value are deleted first
  - Successive deletions occur for earlier files, down to files with modification time stamps older than or equal to 10 minutes, or until file system utilization is less than 75 percent.
  - The renamed alert.log files and ms-odl generation files that are over 5 MB, and older than the successively-shorter age intervals are also deleted.
  - Crash files in the /var/log/oracle/crashfiles directory that are more than one day old can be deleted. If the space pressure is not heavy, then the retention time for crash files is the same as for other files. If there are empty directories under /var/log/oracle/crashfiles, these directories are also deleted.

- For the /opt/oracle file system, the deletion policy is similar to the preceding settings. However, the file threshold is 90 percent, and files are deleted until the file system utilization is less than 85 percent.

- When file system utilization is full, the files controlled by the diagHistoryDays and metricHistoryDays attributes are purged using the same purging policy.

- For the / file system, files in the home directories (cellmonitor and celladmin), /tmp, /var/crash, and /var/spool directories that are over 5 MB and older than one day are deleted.

Every hour, MS deletes eligible alerts from the alert history using the following criteria. Alerts are considered eligible if they are stateless or they are stateful alerts which have been resolved.

- If there are less than 500 alerts, then alerts older than 100 days are deleted.
- If there are between 500 and 999 alerts, then the alerts older than 7 days are deleted.
- If there are 1,000 or more alerts, then all eligible alerts are deleted every minute.

Note:

Any directories or files with SAVE in the name are not deleted.
Related Topics

• ALTER CELL
• DESCRIBE CELL
Administering Oracle ASM on Exadata

- **Overview of Oracle Exadata Database Machine Storage**
  Storage in Oracle Exadata Database Machine consists of servers, cell disks, grid disks, Oracle ASM disk groups, and Oracle ASM failure groups.

- **Administering Oracle ASM Disk Groups Using Oracle Exadata Storage Servers**
  There are basic Oracle ASM tasks needed to use Oracle Exadata Storage Servers.

- **Administering Oracle Exadata Storage Server Grid Disks with Oracle ASM**
  Use the following procedures for managing grid disks used with Oracle ASM.

**Related Topics**

- Maintaining the Hard Disks of Exadata Storage Servers

### 3.1 Overview of Oracle Exadata Database Machine Storage

The following image shows Oracle ASM disk groups created from Oracle Exadata Storage Server grid disks. It represents a typical, but simplified configuration, that can be used as a model for building larger storage grids with additional storage servers and disks.
This example storage grid illustrates the following:

- The storage servers in the grid use an RDMA Network Fabric network to connect to the database servers that have a single-instance database or Oracle Real Application Clusters (Oracle RAC) database installation.
- Each storage server contains multiple physical disks.
- Each cell disk represents a physical disk and a LUN.
- Each cell disk is partitioned into grid disks.
- Oracle ASM disk groups are created using the grid disks.

Oracle ASM failure groups are created to ensure that files are not mirrored on the same storage server, enabling the system to tolerate the failure of a storage server. The number of failure groups equals the number of storage servers. Each failure group is composed of a subset of grid disks in the Oracle ASM disk group that belong to a single storage server.

3.2 Administering Oracle ASM Disk Groups Using Oracle Exadata Storage Servers

There are basic Oracle ASM tasks needed to use Oracle Exadata Storage Servers.
3.2.1 Understanding Oracle ASM Disk Groups for Oracle Exadata Storage Servers

This topic explains Oracle Automatic Storage Management (Oracle ASM) disk groups, and how to create an Oracle ASM disk group for Oracle Exadata System Software using the `CREATE DISKGROUP` SQL command.

Before creating an Oracle ASM disk group, determine which grid disks belong to the Oracle ASM disk group. It is recommended that you choose similar names for the Oracle ASM disk group and its grid disks whenever possible.

The Oracle Exadata Storage Server grid disks are specified with the following pattern:

```
o/cell_IPaddress/griddisk_name
```

In the preceding syntax, `cell_IPaddress` is the IP address of Oracle Exadata Storage Server, and `griddisk_name` is the name of the grid disk.

The cell discovery strings begin with the `o/` prefix.

When specifying the grid disks to be added to the disk group, consider the following:

- The default Oracle ASM disk name is the grid disk name. Oracle recommends using the default name.
- The default failure group name is the cell name. Oracle recommends using the default name.
Wildcards in the form of '*' may be used in the cell_IPaddress or griddisk_name to add multiple disks with one command. For example:

```
CREATE DISKGROUP reco HIGH REDUNDANCYDISK 'o/*/DATA'*
```

When a failure group is not specified, Oracle ASM adds each disk within its own failure group. However, when the disks are stored on Oracle Exadata Storage Servers and a failure group is not specified, Oracle ASM adds a disk to the failure group for that cell. The failure group name is the cell name.

**Note:**

If a cell is renamed, and a disk from that cell is added to an existing disk group that has disks from that cell, then Oracle ASM adds the new disk to a failure group using the new cell name. To ensure all the disks from the cell are in one failure group, add the disk to the disk group and specify the original failure group name.

To enable Smart Scan predicate offload processing, all disks in a disk group must be Oracle Exadata Storage Server grid disks. You cannot include conventional disks with Oracle Exadata Storage Server grid disks.

- **About Fast Disk Scan Rates**
- **Setting the Oracle ASM Content Type**
  Setting the content.type disk group attribute enhances fault tolerance, especially for normal redundancy disk groups.
- **Creating Oracle ASM Disk Groups**
  You can create Oracle ASM disk groups on Oracle Exadata Storage Server grid disks.

**Related Topics**

- **Naming Conventions for Oracle Exadata Storage Server Grid Disks**
  Using a consistent naming convention helps to identify Exadata components.
- **Oracle ASM Disk Groups**
  An Oracle Automatic Storage Management (Oracle ASM) disk group is the primary storage abstraction within Oracle ASM, and is composed of one or more grid disks.

### 3.2.1.1 About Fast Disk Scan Rates

To achieve fast disk scan rates, it is important to lay out segments with at least 4 MB of contiguous space. This allows disk scans to read 4 MB of data before performing another seek at a different location on disk. To ensure segments are laid out with 4 MB of contiguous space, set the Oracle ASM allocation unit size to 4 MB, and ensure data file extents are also at least 4 MB. The allocation unit can be set with the disk group attribute `AU_SIZE` when creating the disk group.

The following SQL command creates a disk group with the allocation unit set to 4 MB. The compatible.rdbms attribute is set to 11.2.0.2 in order to support both release 11.2.0.2 and release 11.2.0.3 databases in a consolidated environment.
CREATE DISKGROUP data NORMAL REDUNDANCY
DISK 'o/*/data_CD*'
ATTRIBUTE 'compatible.rdbms' = '11.2.0.2',
'compatible.asm' = '11.2.0.3',
'content.type' = 'data',
'cell.smart_scan_capable' = 'TRUE',
'au_size' = '4M';

Related Topics

• Oracle Database SQL Language Reference

3.2.1.2 Setting the Oracle ASM Content Type

Setting the content.type disk group attribute enhances fault tolerance, especially for normal redundancy disk groups.

Commencing with Oracle Grid Infrastructure release 11.2.0.3, Oracle ASM provides administrators with the option to specify the content type associated with each disk group. This capability is provided by the content.type disk group attribute. Three possible settings are allowed: data, recovery, or system. Each content type setting modifies the adjacency measure used by the secondary extent placement algorithm.

The result is that the contents of disk groups with different content type settings are distributed differently across the available disks. This decreases the likelihood that a double failure will result in data loss across multiple normal redundancy disk groups with different content type settings. Likewise, a triple failure is less likely to result in data loss on multiple high redundancy disk groups with different content type settings.

The value of content.type attribute should be set as follows:

• DATA and SPARSE disk groups — data
• RECO disk group — recovery
• DBFS_DG disk group (if present) — system

Following this recommendation enhances fault-tolerance. For example, even if the DATA and RECO disk groups use normal redundancy, at least one of the disk groups will survive if two disks fail simultaneously. Therefore, even if DATA is dismounted, databases can typically be recovered from backup objects in RECO.

Note:

• Do not use the content.type attribute to distinguish the availability characteristics of disk groups that are used for a different purpose, such as those created to support a particular service.
• The Oracle Database and Oracle Grid Infrastructure software must be release 12.1.0.2.0 BP5 or later when using sparse grid disks.

1. Use the ALTER DISKGROUP command to set the content.type attribute for an existing disk group, and then rebalance the disk group.
For example:

```sql
ALTER DISKGROUP reco SET ATTRIBUTE 'content.type'='recovery';
ALTER DISKGROUP reco REBALANCE POWER preferred_power_setting;
```

The rebalance operation can take a long time, but the data in the disk group is fully redundant throughout the operation. Oracle ASM monitors the rebalance operation, and Oracle Exadata System Software sends an e-mail message when the operation completes.

2. Check the `content.type` attributes using the following query:

```sql
SQL> SELECT dg.name,a.value FROM v$asm_diskgroup dg, v$asm_attribute a WHERE dg.group_number=a.group_number 
AND a.name='content.type' AND (dg.name LIKE 'DATA%' 
OR dg.name LIKE 'RECO%' OR dg.name LIKE 'DBFS_DG%');
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA</td>
<td>data</td>
</tr>
<tr>
<td>RECO</td>
<td>recovery</td>
</tr>
<tr>
<td>DBFS_DG</td>
<td>system</td>
</tr>
</tbody>
</table>

**Example 3-1 Specifying content.type While Creating a Disk Group**

In this example, the `compatible.rdbms` attribute is set to 11.2.0.2 in order to support both Oracle Database release 11.2.0.2 and release 11.2.0.3 databases in a consolidated environment.

```sql
CREATE DISKGROUP data NORMAL REDUNDANCY 
DISK '/o/*/DATA*' 
ATTRIBUTE 'content.type' = 'DATA', 
'AU_SIZE' = '4M', 
'cell.smart_scan_capable'='TRUE', 
'compatible.rdbms'= '11.2.0.2', 
'compatible.asm'= '11.2.0.3';
```

### 3.2.1.3 Creating Oracle ASM Disk Groups

You can create Oracle ASM disk groups on Oracle Exadata Storage Server grid disks. To create an Oracle ASM disk group to use Oracle Exadata Storage Server grid disks, perform the following procedure:

1. Connect to the Oracle ASM instance.
2. Ensure that the `ORACLE_SID` environment variable is set to the Oracle ASM instance using a command similar to the following:

```bash
$ setenv ORACLE_SID ASM_instance_SID
```
3. Start SQL*Plus on the Oracle ASM instance, and log in as a user with SYSASM administrative privileges.

   $ sqlplus / AS SYSASM

4. Determine which Oracle Exadata Storage Server grid disks are available by querying the V$ASM_DISK view on the Oracle ASM instance, using the following syntax:

   SQL> SELECT path, header_status STATUS FROM V$ASM_DISK WHERE path LIKE '%o/%';

5. Create an Oracle ASM disk group to include disks on the cells.

   In this example, the ALTER command is needed to change compatible.rdbms for the disk group created during installation to hold the OCR and voting disks. The compatible.rdbms attribute is set to 11.2.0.2 in order to support both Oracle Database release 11.2.0.2 and release 11.2.0.3 databases in a consolidated environment.

   CREATE DISKGROUP data HIGH REDUNDANCY
   DISK 'o/*/DATA*' 
   ATTRIBUTE 'AU_SIZE' = '4M',
   'content.type' = 'data',
   'compatible.rdbms'='11.2.0.4',
   'compatible.asm'='19.0.0.0';

   SQL> CREATE DISKGROUP reco HIGH REDUNDANCY 
   DISK 'o/*/RECO*' 
   ATTRIBUTE 'AU_SIZE' = '4M',
   'content.type' = 'recovery',
   'compatible.rdbms'='11.2.0.4',
   'compatible.asm'='19.0.0.0';

   REM for Exadata systems prior to X7
   SQL> ALTER DISKGROUP dbfs_dg SET ATTRIBUTE 
   'content.type' = 'system',
   'compatible.rdbms' = '11.2.0.4';

   When creating sparse disk groups, use a command similar to the following:

   SQL> CREATE DISKGROUP sparsedg NORMAL REDUNDANCY 
   DISK 'o.*/sparse_*' 
   ATTRIBUTE 'AU_SIZE' = '4M',
   'content.type' = 'data',
   'cell.smart_scan_capable'='TRUE',
   'compatible.rdbms' = '12.1.0.2',
   'compatible.asm' = '19.0.0.0',
   'cell.sparse_dg' = 'allsparse';

   In the preceding command, the cell.sparse_dg attribute defines the disk group as a sparse disk group. The attribute is not required if the disk group is not a sparse disk group.
6. View the Oracle ASM disk groups and associated attributes with a SQL query on V$ASM dynamic views.

```
SQL> SELECT dg.name AS diskgroup, SUBSTR(a.name,1,24) AS name, SUBSTR(a.value,1,24) AS value FROM V$ASM_DISKGROUP dg, V$ASM_ATTRIBUTE a WHERE dg.group_number = a.group_number;
```

<table>
<thead>
<tr>
<th>DISKGROUP</th>
<th>NAME</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA</td>
<td>compatible.rdbms</td>
<td>11.2.0.4</td>
</tr>
<tr>
<td>DATA</td>
<td>compatible.asm</td>
<td>19.0.0.0</td>
</tr>
<tr>
<td>DATA</td>
<td>au_size</td>
<td>4194304</td>
</tr>
<tr>
<td>DATA</td>
<td>disk_repair_time</td>
<td>3.6h</td>
</tr>
<tr>
<td>DATA</td>
<td>cell.smart_scan_capable</td>
<td>TRUE</td>
</tr>
</tbody>
</table>
...

7. Create a tablespace in the disk group to take advantage of Oracle Exadata System Software features, such as offload processing. The tablespace should contain the tables that you want to query with offload processing.

```
SQL> CREATE TABLESPACE tablespace_name DATAFILE '+DATA';
```

In the preceding command, +DATA is the name of the Oracle ASM disk group.

8. Verify that the tablespace is in an Oracle Exadata Storage Server disk group. The PREDICATE_EVALUATION column of the DBA_TABLESPACES view indicates whether predicates are evaluated by host (HOST) or by storage (STORAGE).

```
SQL> SELECT tablespace_name, predicate_evaluation FROM dba_tablespaces WHERE tablespace_name = 'DATA_TB';
```

<table>
<thead>
<tr>
<th>TABLESPACE_NAME</th>
<th>PREDICATE_EVALUATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA_TB</td>
<td>STORAGE</td>
</tr>
</tbody>
</table>

**Related Topics**

- Introducing Oracle Automatic Storage Management
3.2.2 Adding a Disk to an Oracle ASM Disk Group

You can add a disk to an Oracle ASM disk group. You might need to do this if you are adding a new Oracle Exadata Storage Server or managing a custom disk group.

Do not add Oracle Exadata Storage Server grid disks to an Oracle ASM disk group that is not on an Oracle Exadata Storage Server unless you are planning to migrate the disk group to an Oracle Exadata Storage Server disk group.

1. Determine which disks are available by querying the `V$ASM_DISK` view on the Oracle ASM instance. If the header status is set to `CANDIDATE`, then the disk is a candidate for a disk group.

2. Use the SQL command `ALTER DISKGROUP` with the `ADD DISK` clause to add the disk to the Oracle ASM disk group.

   For example:

   ```
   SQL> ALTER DISKGROUP disk_group_name
   ADD DISK 'o/cell_IP_address/grid_disk_prefix*';
   ```

   After the disk is added, Oracle ASM rebalances the disk group. Oracle ASM monitors the rebalance operation, and Oracle Exadata System Software sends an e-mail message when the operation is complete.

   You can query the `V$ASM_OPERATION` view for the status of the rebalance operation.

   Related Topics

   • Naming Conventions for Oracle Exadata Storage Server Grid Disks
     Using a consistent naming convention helps to identify Exadata components.

3.2.3 Mounting or Dismounting an Oracle ASM Disk Group

A disk group must be mounted by Oracle ASM before a database can access the files in the disk group.

Mounting a disk group requires discovering all of the disks and locating the files in the disk group. When an Oracle ASM starts, the disk groups mentioned in the `ASM_DISKGROUPS` instance parameter are automatically mounted.

Additionally:

• To mount a disk group, you can use the SQL `ALTER DISKGROUP` command with the `MOUNT` option.
You can use the `FORCE` option in conjunction with the `ALTER DISKGROUP ... MOUNT` command to mount a disk group even if a disk is unavailable. However, this compromises redundancy in the disk group.

- To dismount a disk group, you can use the SQL `ALTER DISKGROUP` command with the `DISMOUNT` option.

Related Topics
- *Oracle Automatic Storage Management Administrator's Guide*

### 3.2.4 Changing a Disk to Offline or Online

You can change an Oracle ASM disk to **INACTIVE** or **ACTIVE**.

1. Determine which disk you want offline or online in the Oracle ASM disk group. Query the `V$ASM_DISK` and `V$ASM_DISKGROUP` views on the Oracle ASM instance.

2. Use one of the following commands:
   - To make a disk inactive, use the following command:
     ```
     CellCLI> ALTER GRIDDISK gdisk_name INACTIVE
     ```
   - To make a disk active, use the following command:
     ```
     CellCLI> ALTER GRIDDISK gdisk_name ACTIVE
     ```

As soon as the disk is online, the disk group is rebalanced. Oracle ASM monitors the rebalance operation, and Oracle Exadata System Software sends an e-mail message when the operation is complete.

You can query the `V$ASM_OPERATION` view for the status of the rebalance operation.

Related Topics
- **Determining Which Oracle ASM Disk Group Contains an Oracle Exadata Storage Server Grid Disk**
  If a grid disk name matches the Oracle ASM disk name, and the name contains the Oracle ASM disk group name, then you can determine the Oracle ASM disk group to which the grid disk belongs.

- **ALTER GRIDDISK**

### 3.2.5 Dropping a Disk from an Oracle ASM Disk Group

You can drop a grid disk from a disk group.

1. Determine which disks you want to drop from the Oracle ASM disk group. Query the `V$ASM_DISK` and `V$ASM_DISKGROUP` views on the Oracle ASM instance.

   If you are removing an Oracle Exadata Storage Server grid disk, then ensure that you identify the grid disks that are mapped to each Oracle ASM disk group.
2. Use the SQL ALTER DISKGROUP command with the DROP DISK clause to drop the disks from the Oracle ASM disk group.

    SQL> ALTER DISKGROUP disk_group_name
        DROP DISK data_CD_11_cell01 NORMAL;

Do not use the FORCE option when dropping the disk from the Oracle ASM disk group. If you use the FORCE option, Oracle Exadata System Software will attempt to add the disk back to the disk group if the disk online automation operation is triggered, by rebooting the storage server, for example. See Enhanced Manageability Features in Oracle Exadata Database Machine System Overview.

When the disk is dropped from the Oracle ASM disk group, Oracle ASM rebalances the disk group. Oracle ASM monitors the rebalance operation, and Oracle Exadata System Software sends an e-mail message when the operation is complete.

You can query the V$ASM_OPERATION view for the status of the rebalance operation.

After an Oracle Exadata Storage Server grid disk is dropped from the Oracle ASM disk group, you can drop the grid disk from the cell.

Related Topics

- Determining Which Oracle ASM Disk Group Contains an Oracle Exadata Storage Server Grid Disk
  If a grid disk name matches the Oracle ASM disk name, and the name contains the Oracle ASM disk group name, then you can determine the Oracle ASM disk group to which the grid disk belongs.

- Determining Which Oracle Exadata Storage Server Grid Disks Belong to an Oracle ASM Disk Group
  If a grid disk name contains the Oracle ASM disk group name, then you can use SQL commands on the Oracle ASM instance to list the Oracle ASM disk group names.

- Dropping an Oracle Exadata Storage Server Grid Disk
  To drop an Oracle Exadata Storage Server grid disk, use the CellCLI DROP GRIDDISK command.

### 3.2.6 Dropping an Oracle ASM Disk Group

You can drop an Oracle ASM disk group.

If you cannot mount a disk group but must drop it, then use the FORCE option with the DROP DISKGROUP command.

1. Determine the disk group that you want to drop.
   Query the V$ASM_DISKGROUP view on the Oracle ASM instance.

2. Use the SQL DROP DISKGROUP command to drop the Oracle ASM disk group.
3.2.7 Enabling the Oracle ASM appliance.mode Attribute

The Oracle ASM appliance.mode attribute improves disk rebalance completion time when dropping one or more Oracle ASM disks.

Setting the appliance.mode attribute helps restore redundancy faster after a failure. The attribute can only be enabled on disk groups that meet the following requirements:

- The Oracle ASM disk group attribute compatible.asm is set to release 11.2.0.4, or 12.1.0.2 or later.
- The cell.smart_scan_capable attribute is set to TRUE.
- All disks in the disk group are the same type; for example, all disks are hard disks or all disks are flash disks.
- All disks in the disk group are the same size.
- All failure groups in the disk group have an equal number of disks:
  - For eighth rack configurations, all failure groups have 4 disks, or all failure groups have 6 disks.
  - For all other rack configurations, all failure groups have 10 disks, or all failure groups have 12 disks.
- There are at least 3 failure groups in the disk group.
- No disk in the disk group is offline.

**Note:**

Enabling the appliance.mode attribute for existing disk groups may cause an increase of data movement during the next rebalance operation.

The appliance.mode attribute is automatically enabled when creating a new disk group. Existing disk groups must explicitly set the attribute using the ALTER DISKGROUP command.

```sql
SQL> ALTER DISKGROUP disk_group SET ATTRIBUTE 'appliance.mode'='TRUE';
```

**Note:**

The appliance.mode attribute should normally be set to TRUE. In rare cases it may be necessary to disable appliance.mode as a workaround when adding disks to a disk group. After the disk group is ALTERed enable appliance.mode, and perform a REBALANCE operation.
To disable the `appliance.mode` attribute during disk group creation, set the attribute to `FALSE`.

```
SQL> CREATE DISKGROUP data NORMAL REDUNDANCY
       DISK
       'o/*'/DATA*
       ATTRIBUTE 'content.type' = 'data',
               'au_size' = '4M',
               'cell.smart_scan_capable'='TRUE',
               'compatible.rdbms'='11.2.0.3',
               'compatible.asm'='11.2.0.4',
               'appliance.mode'='FALSE';
```

3.2.8 Checking Disk Group Balance

Files should be equally balanced across all disks. The following queries and script can be used to check disk group balance:

- To check I/O balance, query the `V$ASM_DISK_IOSTAT` view before and after running a large SQL statement. For example, if a large query has a lot of reads, then the values in the `read` column and the `read_bytes` column should be approximately the same for all disks in the disk group.
- To check all mounted disk groups, run the script available in My Oracle Support document 367445.1.

3.2.9 Setting the Oracle ASM Disk Repair Timer

The Oracle ASM disk repair timer represents the amount of time a disk can remain offline before it is dropped by Oracle ASM. While the disk is offline, Oracle ASM tracks the changed extents so the disk can be resynchronized when it comes back online. The default disk repair time is 3.6 hours. If the default is inadequate, then the attribute value can be changed to the maximum amount of time it might take to detect and repair a temporary disk failure. The following command is an example of changing the disk repair timer value to 8.5 hours for the `DATA` disk group:

```
SQL> ALTER DISKGROUP data SET ATTRIBUTE 'disk_repair_time' = '8.5h'
```

The `disk_repair_time` attribute does not change the repair timer for disks currently offline. The repair timer for those offline disks is either the default repair timer or the repair timer specified on the command line when the disks were manually set to offline. To change the repair timer for currently offline disks, use the `OFFLINE` command and specify a repair timer value. The following command is an example of changing the disk repair timer value for disks that are offline:

```
SQL> ALTER DISKGROUP data OFFLINE DISK data_CD_06_cell11 DROP AFTER 20h;
```
Note:
Vulnerability to a double failure increases in line with increases to the disk repair time value.

Related Topics
• Oracle Automatic Storage Management Administrator's Guide

3.3 Administering Oracle Exadata Storage Server Grid Disks with Oracle ASM

Use the following procedures for managing grid disks used with Oracle ASM.

• Naming Conventions for Oracle Exadata Storage Server Grid Disks
Using a consistent naming convention helps to identify Exadata components.

• Changing an Oracle Exadata Storage Server Grid Disk That Belongs to an Oracle ASM Disk Group
Before you change a grid disk that belongs to an Oracle ASM disk group, you must consider how the change might affect the Oracle ASM disk group to which the grid disk belongs.

• Resizing Grid Disks
You can resize grid disks and Oracle ASM disk groups to shrink one with excess free space and increase the size of another that is near capacity.

• Determining Which Oracle ASM Disk Group Contains an Oracle Exadata Storage Server Grid Disk
If a grid disk name matches the Oracle ASM disk name, and the name contains the Oracle ASM disk group name, then you can determine the Oracle ASM disk group to which the grid disk belongs.

• Determining Which Oracle Exadata Storage Server Grid Disks Belong to an Oracle ASM Disk Group
If a grid disk name contains the Oracle ASM disk group name, then you can use SQL commands on the Oracle ASM instance to list the Oracle ASM disk group names.

• Handling Disk Replacement
If a disk has a problem, the physical disk status changes.

3.3.1 Naming Conventions for Oracle Exadata Storage Server Grid Disks

Using a consistent naming convention helps to identify Exadata components.

The name of the grid disk should contain the cell disk name to make it easier to determine which grid disks belong to a cell disk. To help determine which grid disks belong to an Oracle ASM disk group, a subset of the grid disk name should match all or part of the name of the Oracle ASM disk group to which the grid disk will belong.
For example, if a grid disk is created on the cell disk CD_03_cell01, and that grid disk belongs to an Oracle ASM disk group named data0, then the grid disk name should be data0_CD_03_cell01.

When you use the **ALL PREFIX** option with **CREATE GRIDDISK**, a unique grid disk name is automatically generated that includes the prefix and cell name. If you do not use the default generated name when creating grid disks, then you must ensure that the grid disk name is unique across all cells. You cannot have multiple disks with the same name in an Oracle ASM disk group.

**Related Topics**

- **Determining Which Oracle ASM Disk Group Contains an Oracle Exadata Storage Server Grid Disk**
  If a grid disk name matches the Oracle ASM disk name, and the name contains the Oracle ASM disk group name, then you can determine the Oracle ASM disk group to which the grid disk belongs.

- **Determining Which Oracle Exadata Storage Server Grid Disks Belong to an Oracle ASM Disk Group**
  If a grid disk name contains the Oracle ASM disk group name, then you can use SQL commands on the Oracle ASM instance to list the Oracle ASM disk group names.

- **CREATE CELLDISK**
- **CREATE GRIDDISK**

### 3.3.2 Changing an Oracle Exadata Storage Server Grid Disk That Belongs to an Oracle ASM Disk Group

Before you change a grid disk that belongs to an Oracle ASM disk group, you must consider how the change might affect the Oracle ASM disk group to which the grid disk belongs.

- **Changing an Oracle Exadata Storage Server Grid Disk Name**
  Use the CellCLI interface to change the name of a grid disk.

- **Dropping an Oracle Exadata Storage Server Grid Disk**
  To drop an Oracle Exadata Storage Server grid disk, use the CellCLI **DROP GRIDDISK** command.

#### 3.3.2.1 Changing an Oracle Exadata Storage Server Grid Disk Name

Use the CellCLI interface to change the name of a grid disk.

- **To change attributes of a grid disk, use the CellCLI **ALTER GRIDDISK** command.**

  Use the **DESCRIBE GRIDDISK** command to determine which Oracle Exadata Storage Server grid disk attributes can be modified.

**Caution:**

Before changing the name of a grid disk that belongs to an Oracle ASM disk group, ensure that the corresponding Oracle ASM disk is offline.
Example 3-2  Changing an Oracle Exadata Storage Server Grid Disk Name

Use the ALTER GRIDDISK command to rename a grid disk.

CellCLI> ALTER GRIDDISK data011 name='data0_CD_03_cell04'

Related Topics

• Changing a Disk to Offline or Online
  You can change an Oracle ASM disk to INACTIVE or ACTIVE.

• ALTER GRIDDISK

• Naming Conventions for Oracle Exadata Storage Server Grid Disks
  Using a consistent naming convention helps to identify Exadata components.

3.3.2.2 Dropping an Oracle Exadata Storage Server Grid Disk

To drop an Oracle Exadata Storage Server grid disk, use the CellCLI DROP GRIDDISK command.

Make the grid disk inactive before dropping the grid disk to ensure that the grid disk is not in use. The FORCE option can be used to force the grid disk that is in use to be dropped.

Caution:

• Before dropping a grid disk that belongs to an Oracle ASM disk group, ensure that the corresponding Oracle ASM disk was dropped from the disk group.

• Before dropping a grid disk using the FORCE option, ensure that the Oracle ASM disk was dropped from the disk group. If you drop a grid disk that is still part of an ASM disk group, you may compromise data redundancy in the disk group or cause the disk group to dismount.

1. Drop the Oracle ASM disk from the disk group.

   SQL> ALTER DISKGROUP disk_group_name DROP DISK disk_name;

2. Make the corresponding grid disk inactive.

   CellCLI> ALTER GRIDDISK disk_name INACTIVE

3. Drop the grid disk.

   CellCLI> DROP GRIDDISK disk_name
Example 3-3  Dropping a specific grid disk

After you have dropped the Oracle ASM disk from the disk group, you can drop the related grid disk.

CellCLI> ALTER GRIDDISK data0_CD_03_cell04 INACTIVE
CellCLI> DROP GRIDDISK data0_CD_03_cell04

Example 3-4  Dropping all grid disks

After you have dropped the Oracle ASM disks from the disk group, you can drop multiple grid disks using a single command.

CellCLI> ALTER GRIDDISK ALL INACTIVE
CellCLI> DROP GRIDDISK ALL PREFIX=data0

Example 3-5  Using the FORCE option when dropping a grid disk

The FORCE option forces an active grid disk to be dropped. For example, if you cannot make a grid disk INACTIVE, but must drop the grid disk, you can use the FORCE option.

Use the FORCE option cautiously. If you drop a grid disk that is still part of an ASM disk group, you may compromise data redundancy in the disk group or cause the disk group to dismount.

CellCLI> DROP GRIDDISK data02_CD_04_cell01 FORCE

Related Topics

- DROP GRIDDISK

3.3.3 Resizing Grid Disks

You can resize grid disks and Oracle ASM disk groups to shrink one with excess free space and increase the size of another that is near capacity.

Initial configuration of Oracle Exadata Database Machine disk group sizes is based on Oracle best practices and the location of the backup files.

- For internal backups: allocation of available space is 40% for the DATA disk groups, and 60% for the RECO disk groups.
- For external backups: allocation of available space is 80% for the DATA disk group, and 20% for the RECO disk group.

The disk group allocations can be changed after deployment. For example, the DATA disk group allocation may be too small at 60%, and need to be resized to 80%.

If your system has no free space available on the cell disks and one disk group, for example RECO, has plenty of free space, then you can resize the RECO disk group to a smaller size and reallocate the free space to the DATA disk group. The free space available after shrinking the RECO disk group is at a non-contiguous offset from the existing space allocations for the DATA disk group. Grid disks can use space anywhere on the cell disks and do not have to be contiguous.

If you are expanding the grid disks and the cell disks already have sufficient space to expand the existing grid disks, then you do not need to first resize an existing disk.
group. You would skip steps 2 and 3 below where the example shows the RECO disk group and grid disks are shrunk (you should still verify the cell disks have enough free space before growing the DATA grid disks). The amount of free space the administrator should reserve depends on the level of failure coverage.

If you are shrinking the size of the grid disks, you should understand how space is reserved for mirroring. Data is protected by Oracle ASM using normal or high redundancy to create one or two copies of data, which are stored as file extents. These copies are stored in separate failure groups. A failure in one failure group does not affect the mirror copies, so data is still accessible.

When a failure occurs, Oracle ASM re-mirrors, or rebalances, any extents that are not accessible so that redundancy is reestablished. For the re-mirroring process to succeed, sufficient free space must exist in the disk group to allow creation of the new file extent mirror copies. If there is not enough free space, then some extents will not be re-mirrored and the subsequent failure of the other data copies will require the disk group to be restored from backup. Oracle ASM sends an error when a re-mirror process fails due to lack of space.

You must be using Oracle Exadata System Software release 12.1.2.1.0 or higher, or have the patch for bug 19695225 applied to your software.

This procedure for resizing grid disks applies to bare metal and virtual machine (VM) deployments.

• Determine the Amount of Available Space
  To increase the size of the disks in a disk group you must either have unallocated disk space available, or you have to reallocate space currently used by a different disk group.

• Shrink the Oracle ASM Disks in the Donor Disk Group
  If there is no free space available on the cell disks, you can reduce the space used by one disk group to provide additional disk space for a different disk group.

• Shrink the Grid Disks in the Donor Disk Group
  After shrinking the disks in the Oracle ASM disk group, you then shrink the size of the grid disks on each cell.

• Increase the Size of the Grid Disks Using Available Space
  You can increase the size used by the grid disks if there is unallocated disk space either already available, or made available by shrinking the space used by a different Oracle ASM disk group.

• Increase the Size of the Oracle ASM Disks
  You can increase the size used by the Oracle ASM disks after increasing the space allocated to the associated grid disks.

Related Topics
• Understanding ASM Capacity and Reservation of Free Space in Exadata (My Oracle Support Doc ID 1551288.1)
• Bug 19695225 - Running Many Create or Alter Griddisk Commands Over Time Causes Cell Disk Metadata Corruption (ORA-600 [addNewSegmentsToGDisk_2]) and Loss of Cell Disk Content (My Oracle Support Doc ID 1991445.1)
3.3.3.1 Determine the Amount of Available Space

To increase the size of the disks in a disk group you must either have unallocated disk space available, or you have to reallocate space currently used by a different disk group.

You can also use a script available in "Script to Calculate New Grid Disk and Disk Group Sizes in Exadata (My Oracle Support Doc ID 1464809.1)" to assist in determining how much free space is available to shrink a disk group.

1. View the space currently used by the disk groups.

   ```sql
   SELECT name, total_mb, free_mb, total_mb - free_mb used_mb,
          round(100*free_mb/total_mb,2) pct_free
   FROM v$asm_diskgroup
   ORDER BY 1;
   ```

<table>
<thead>
<tr>
<th>NAME</th>
<th>TOTAL_MB</th>
<th>FREE_MB</th>
<th>USED_MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>68812800</td>
<td>9985076</td>
<td>58827724</td>
</tr>
<tr>
<td>RECOC1</td>
<td>94980480</td>
<td>82594920</td>
<td>12385560</td>
</tr>
</tbody>
</table>

   The example above shows that the DATAC1 disk group has only about 15% of free space available while the RECOC1 disk group has about 87% free disk space. The PCT_FREE displayed here is raw free space, not usable free space. Additional space is needed for rebalancing operations.

2. For the disk groups you plan to resize, view the count and status of the failure groups used by the disk groups.

   ```sql
   SELECT dg.name, d.failgroup, d.state, d.header_status,
          d.mount_mode,
          d.mode_status, count(1) num_disks
   FROM v$asm_disk d, v$asm_diskgroup dg
   WHERE d.group_number = dg.group_number
   AND dg.name IN ('RECOC1', 'DATAC1')
   GROUP BY dg.name, d.failgroup, d.state, d.header_status,
            d.mount_status,
            d.mode_status
   ORDER BY 1, 2, 3;
   ```

<table>
<thead>
<tr>
<th>NAME</th>
<th>FAILGROUP</th>
<th>STATE</th>
<th>HEADER_STATUS</th>
<th>MOUNT_Mode</th>
<th>MODE_STATUS</th>
<th>NUM_DISKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
<td>12</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM02</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
<td>12</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM03</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM04</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM05</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM06</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM07</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM08</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM09</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM10</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM11</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM12</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>DATAC1</td>
<td>EXA01CELADM13</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM01</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM02</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM03</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM04</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM05</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM06</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM07</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM08</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM09</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM10</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM11</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM12</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM13</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
<tr>
<td>12</td>
<td>RECOC1</td>
<td>EXA01CELADM14</td>
<td>NORMAL</td>
<td>MEMBER</td>
<td>CACHED</td>
<td>ONLINE</td>
</tr>
</tbody>
</table>

The above example is for a full rack, which has 14 cells and 14 failure groups for DATAC1 and RECOC1. Verify that each failure group has at least 12 disks in
the NORMAL state (num_disks). If you see disks listed as MISSING, or you see an unexpected number of disks for your configuration, then do not proceed until you resolve the problem.

Extreme Flash systems should see a disk count of 8 instead of 12 for num_disks.

3. List the corresponding grid disks associated with each cell and each failure group, so you know which grid disks to resize.

```
SELECT dg.name, d.failgroup, d.path
FROM V$ASM_DISK d, V$ASM_DISKGROUP dg
WHERE d.group_number = dg.group_number
AND dg.name IN ('RECOC1','DATAC1')
ORDER BY 1, 2, 3;
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>FAILGROUP</th>
<th>PATH</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_00</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_01</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_02</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_03</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_04</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_05</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_06</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_07</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_08</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_09</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_10</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1_CD_11</td>
<td>EXA01CELADM01</td>
<td>o/192.168.74.43/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM02</td>
<td>o/192.168.74.44/</td>
</tr>
<tr>
<td>DATAC1_CD_00</td>
<td>EXA01CELADM02</td>
<td>o/192.168.74.44/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM02</td>
<td>o/192.168.74.44/</td>
</tr>
<tr>
<td>DATAC1_CD_01</td>
<td>EXA01CELADM02</td>
<td>o/192.168.74.44/</td>
</tr>
<tr>
<td>DATAC1</td>
<td>EXA01CELADM02</td>
<td>o/192.168.74.44/</td>
</tr>
<tr>
<td>DATAC1_CD_02</td>
<td>EXA01CELADM02</td>
<td>o/192.168.74.44/</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RECOC1</td>
<td>EXA01CELADM13</td>
<td>o/192.168.74.55/</td>
</tr>
<tr>
<td>RECOC1_CD_00</td>
<td>EXA01CELADM13</td>
<td>o/192.168.74.55/</td>
</tr>
<tr>
<td>RECOC1</td>
<td>EXA01CELADM13</td>
<td>o/192.168.74.55/</td>
</tr>
<tr>
<td>RECOC1_CD_01</td>
<td>EXA01CELADM13</td>
<td>o/192.168.74.55/</td>
</tr>
<tr>
<td>RECOC1</td>
<td>EXA01CELADM13</td>
<td>o/192.168.74.55/</td>
</tr>
<tr>
<td>RECOC1_CD_02</td>
<td>EXA01CELADM13</td>
<td>o/192.168.74.55/</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. Check the cell disks for available free space.

Free space on the cell disks can be used to increase the size of the DATAC1 grid disks. If there is not enough available free space to expand the DATAC1 grid disks, then you must shrink the RECOC1 grid disks to provide the additional space for the desired new size of DATAC1 grid disks.

```
[root@exa01adm01 tmp]# dcli -g ~/cell_group -l root "cellcli -e
list celldisk \
    attributes name,freespace"
exa01celadm01: CD_00_exa01celadm01 0
exa01celadm01: CD_01_exa01celadm01 0
exa01celadm01: CD_02_exa01celadm01 0
exa01celadm01: CD_03_exa01celadm01 0
exa01celadm01: CD_04_exa01celadm01 0
exa01celadm01: CD_05_exa01celadm01 0
exa01celadm01: CD_06_exa01celadm01 0
exa01celadm01: CD_07_exa01celadm01 0
exa01celadm01: CD_08_exa01celadm01 0
exa01celadm01: CD_09_exa01celadm01 0
exa01celadm01: CD_10_exa01celadm01 0
exa01celadm01: CD_11_exa01celadm01 0
...
```

In this example, there is no free space available, so you must shrink the RECOC1 grid disks first to provide space for the DATAC1 grid disks. In your configuration there might be plenty of free space available and you can use that free space instead of shrinking the RECOC1 grid disks.

5. Calculate the amount of space to shrink from the RECOC1 disk group and from each grid disk.

The minimum size to safely shrink a disk group and its grid disks must take into account the following:

- Space currently in use (**USED_MB**)
- Space expected for growth (**GROWTH_MB**)
- Space needed to rebalance in case of disk failure (**DFC_MB**), typically 15% of total disk group size

The minimum size calculation taking the above factors into account is:

\[
\text{Minimum DG size (MB)} = (\text{USED_MB} + \text{GROWTH_MB}) \times 1.15
\]

**USED_MB** can be derived from V$ASM_DISKGROUP by calculating **TOTAL_MB** - **FREE_MB**
• GROWTH_MB is an estimate specific to how the disk group will be used in the future and should be based on historical patterns of growth.

For the RECO1 disk group space usage shown in step 1, we see the minimum size it can shrink to assuming no growth estimates is:

Minimum RECO1 size = (TOTAL_MB - FREE_MB + GROWTH_MB) * 1.15

= ( 94980480 - 82594920 + 0 ) * 1.15 = 14243394 MB = 13,910 GB

In the example output shown in Step 1, RECO1 has plenty of free space and DATAC1 has less than 15% free. So, you could shrink RECO1 and give the freed disk space to DATAC1. If you decide to reduce RECO1 to half of its current size, the new size is 94980480 / 2 = 47490240 MB. This size is significantly above the minimum size we calculated for the RECO1 disk group above, so it is safe to shrink it down to this value.

The query in Step 2 shows that there are 168 grid disks for RECO1, because there are 14 cells and 12 disks per cell (14 * 12 = 168). The estimated new size of each grid disk for the RECO1 disk group is 47490240 / 168, or 282,680 MB.

Find the closest 16 MB boundary for the new grid disk size. If you do not perform this check, then the cell will round down the grid disk size to the nearest 16 MB boundary automatically, and you could end up with a mismatch in size between the Oracle ASM disks and the grid disks.

SQL> SELECT 16*TRUNC(new_disk_size/16) new_disk_size FROM dual;
Enter value for new_disk_size: 282680

<table>
<thead>
<tr>
<th>NEW_DISK_SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>282672</td>
</tr>
</tbody>
</table>

Based on the above result, you should choose 282672 MB as the new size for the grid disks in the RECO1 disk group. After resizing the grid disks, the size of the RECO1 disk group will be 47488896 MB.

6. Calculate how much to increase the size of each grid disk in the DATAC1 disk group.

Ensure the Oracle ASM disk size and the grid disk sizes match across the entire disk group. The following query shows the combinations of disk sizes in each disk group. Ideally, there is only one size found for all disks and the sizes of both the Oracle ASM (total_mb) disks and the grid disks (os_mb) match.

```
SELECT dg.name, d.total_mb, d.os_mb, count(1) num_disks
FROM v$asm_diskgroup dg, v$asm_disk d
WHERE dg.group_number = d.group_number
GROUP BY dg.name, d.total_mb, d.os_mb;
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>TOTAL_MB</th>
<th>OS_MB</th>
<th>NUM_DISKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>409600</td>
<td>409600</td>
<td>168</td>
</tr>
<tr>
<td>RECO1</td>
<td>565360</td>
<td>565360</td>
<td>168</td>
</tr>
</tbody>
</table>

After shrinking RECO1's grid disks, the following space is left per disk for DATAC1:
Additional space for DATAC1 disks = \textit{RECOC1\_current\_size} - \textit{RECOC1\_new\_size} = 565360 - 282672 = 282688 MB

To calculate the new size of the grid disks for the DATAC1 disk group, use the following:

\[
\text{DATAC1\_disks\_new\_size} = \text{DATAC1\_disks\_current\_size} + \text{new\_free\_space\_from\_RECOC1}
\]
\[
= 409600 + 282688 = 692288 \text{ MB}
\]

Find the closest 16 MB boundary for the new grid disk size. If you do not perform this check, then the cell will round down the grid disk size to the nearest 16 MB boundary automatically, and you could end up with a mismatch in size between the Oracle ASM disks and the grid disks.

\[
\text{SQL} > \text{SELECT 16*TRUNC(\&new\_disk\_size/16) new\_disk\_size FROM dual;}
\]

\[
\text{Enter value for new\_disk\_size: 692288}
\]

\[
\begin{array}{|c|}
\hline
\text{NEW\_DISK\_SIZE} \\
\hline
\text{692288} \\
\hline
\end{array}
\]

Based on the query result, you can use the calculated size of 692288 MB for the disks in the DATAC1 disk groups because the size is on a 16 MB boundary. If the result of the query is different from the value you supplied, then you must use the value returned by the query because that is the value to which the cell will round the grid disk size.

The calculated value of the new grid disk size will result in the DATAC1 disk group having a total size of 116304384 MB (168 disks * 692288 MB).

3.3.3.2 Shrink the Oracle ASM Disks in the Donor Disk Group

If there is no free space available on the cell disks, you can reduce the space used by one disk group to provide additional disk space for a different disk group.

This task is a continuation of an example where space in the RECOC1 disk group is being reallocated to the DATAC1 disk group.

Before resizing the disk group, make sure the disk group you are taking space from has sufficient free space.

1. Shrink the Oracle ASM disks for the RECO disk group down to the new desired size for all disks.

   Use the new size for the disks in the RECO disk group that was calculated in Step 5 of Determine the Amount of Available Space.

   \[
   \text{SQL} > \text{ALTER DISKGROUP recoc1 resize all size 282672m rebalance power 64;}
   \]
The ALTER DISKGROUP command may take several minutes to complete. The SQL prompt will not return until this operation has completed.  

If the specified disk group has quorum disks configured within the disk group, then the ALTER DISKGROUP ... SIZE ALL command could fail with error ORA-15277. Quorum disks are configured if the requirements specified in Managing Quorum Disks for High Redundancy Disk Groups are met. As a workaround, you can specify the storage server failure group names (for the ones with a FAILURE_TYPE of REGULAR, not QUORUM) explicitly in the SQL command, for example:

```
SQL> ALTER DISKGROUP recoc1 RESIZE DISKS IN FAILGROUP
      exacell01 SIZE 282672M,
      exacell02 SIZE 282672M, exacell03 SIZE 282672M REBALANCE
      POWER 64;
```

Wait for rebalance to finish by checking the view GV$ASM_OPERATION.

```
SQL> set lines 250 pages 1000
SQL> col error_code form a10
SQL> SELECT dg.name, o.*
2      FROM gv$asm_operation o, v$asm_diskgroup dg
3      WHERE o.group_number = dg.group_number;
```

Proceed to the next step ONLY when the query against GV$ASM_OPERATION shows no rows for the disk group being altered.

2. Verify the new size of the ASM disks using the following queries:

```
SQL> SELECT name, total_mb, free_mb, total_mb - free_mb used_mb,
       2      ROUND(100*free_mb/total_mb,2) pct_free
       3      FROM v$asm_diskgroup
       4      ORDER BY 1;
```

```
+-----------------+---------+-------------+---------+
<table>
<thead>
<tr>
<th>NAME</th>
<th>TOTAL_MB</th>
<th>FREE_MB</th>
<th>USED_MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>68812800</td>
<td>9985076</td>
<td>58827724</td>
</tr>
<tr>
<td>RECOC1</td>
<td>47488896</td>
<td>35103336</td>
<td>12385560</td>
</tr>
</tbody>
</table>
```

```
SQL> SELECT dg.name, d.total_mb, d.os_mb, COUNT(1) num_disks
       2      FROM v$asm_diskgroup dg, v$asm_disk d
       3      WHERE dg.group_number = d.group_number
       4      GROUP BY dg.name, d.total_mb, d.os_mb;
```

```
+-----------------+-------+-------+--------+
<table>
<thead>
<tr>
<th>NAME</th>
<th>TOTAL_MB</th>
<th>OS_MB</th>
<th>NUM_DISKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>----------------</td>
<td>---------</td>
<td>-------</td>
<td>----------</td>
</tr>
</tbody>
</table>
```

Chapter 3
Administering Oracle Exadata Storage Server Grid Disks with Oracle ASM
The above query example shows that the disks in the RECOC1 disk group have been resized to a size of 282672 MG each, and the total disk group size is 4748896 MB.

3.3.3.3 Shrink the Grid Disks in the Donor Disk Group

After shrinking the disks in the Oracle ASM disk group, you then shrink the size of the grid disks on each cell.

This task is a continuation of an example where space in the RECOC1 disk group is being reallocated to the DATAC1 disk group.

You must have first completed the task Shrink the Oracle ASM Disks in the Donor Disk Group.

1. Shrinking the grid disks associated with the RECO disk group on all cells down to the new, smaller size.

   For each storage cell identified in Determine the Amount of Available Space in Step 3, shrink the grid disks to match the size of the Oracle ASM disks that were shrunk in the previous task. Use commands similar to the following:

   ```bash
   dcli -c exa01celadm01 -l root "cellcli -e alter griddisk
   RECOC1_CD_00_exa01celadm01 \
   ,RECOC1_CD_01_exa01celadm01 \
   ,RECOC1_CD_02_exa01celadm01 \
   ,RECOC1_CD_03_exa01celadm01 \
   ,RECOC1_CD_04_exa01celadm01 \
   ,RECOC1_CD_05_exa01celadm01 \
   ,RECOC1_CD_06_exa01celadm01 \
   ,RECOC1_CD_07_exa01celadm01 \
   ,RECOC1_CD_08_exa01celadm01 \
   ,RECOC1_CD_09_exa01celadm01 \
   ,RECOC1_CD_10_exa01celadm01 \
   ,RECOC1_CD_11_exa01celadm01 \
   size=282672M "
   ``

   ```bash
   dcli -c exa01celadm02 -l root "cellcli -e alter griddisk
   RECOC1_CD_00_exa01celadm02 \
   ,RECOC1_CD_01_exa01celadm02 \
   ,RECOC1_CD_02_exa01celadm02 \
   ,RECOC1_CD_03_exa01celadm02 \
   ,RECOC1_CD_04_exa01celadm02 \
   ,RECOC1_CD_05_exa01celadm02 \
   ,RECOC1_CD_06_exa01celadm02 \
   ,RECOC1_CD_07_exa01celadm02 \
   ,RECOC1_CD_08_exa01celadm02 \
   ,RECOC1_CD_09_exa01celadm02 \
   ,RECOC1_CD_10_exa01celadm02 \
   ,RECOC1_CD_11_exa01celadm02 \
   size=282672M "
   ```
2. Verify the new size of the grid disks using the following command:

   ```
   [root@exa01adm01 tmp]# dcli -g cell_group -l root "cellcli -e list griddisk attributes name,size where name like '\RECOC1.*'"
   
   exa01celadm01: RECOC1_CD_00_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_01_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_02_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_03_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_04_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_05_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_06_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_07_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_08_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_09_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_10_exa01celadm01 276.046875G
   exa01celadm01: RECOC1_CD_11_exa01celadm01 276.046875G
   ...  
   
   The above example shows that the disks in the RECOC1 disk group have been resized to a size of 282672 MB each (276.046875 * 1024).

3.3.3.4 Increase the Size of the Grid Disks Using Available Space

You can increase the size used by the grid disks if there is unallocated disk space either already available, or made available by shrinking the space used by a different Oracle ASM disk group.

This task is a continuation of an example where space in the RECOC1 disk group is being reallocated to the DATAC1 disk group. If you already have sufficient space to expand an existing disk group, then you do not need to reallocate space from a different disk group.

1. Check that the cell disks have the expected amount of free space.
After completing the tasks to shrink the Oracle ASM disks and the grid disks, you would expect to see the following free space on the cell disks:

```
[root@exa01adm01 tmp]# dcli -g ~/cell_group -l root "cellcli -e list celldisk \ attributes name,freespace"

exa01celadm01:  CD_00_exa01celadm01  276.0625G
exa01celadm01:  CD_01_exa01celadm01  276.0625G
exa01celadm01:  CD_02_exa01celadm01  276.0625G
exa01celadm01:  CD_03_exa01celadm01  276.0625G
exa01celadm01:  CD_04_exa01celadm01  276.0625G
exa01celadm01:  CD_05_exa01celadm01  276.0625G
exa01celadm01:  CD_06_exa01celadm01  276.0625G
exa01celadm01:  CD_07_exa01celadm01  276.0625G
exa01celadm01:  CD_08_exa01celadm01  276.0625G
exa01celadm01:  CD_09_exa01celadm01  276.0625G
exa01celadm01:  CD_10_exa01celadm01  276.0625G
exa01celadm01:  CD_11_exa01celadm01  276.0625G
...
```

2. For each storage cell, increase the size of the DATA grid disks to the desired new size. Use the size calculated in Determine the Amount of Available Space.

```
dcli -c exa01celadm01 -l root "cellcli -e alter griddisk  
  DATA1_CD_00_exa01celadm01 \       
  ,DATA1_CD_01_exa01celadm01 \       
  ,DATA1_CD_02_exa01celadm01 \       
  ,DATA1_CD_03_exa01celadm01 \       
  ,DATA1_CD_04_exa01celadm01 \       
  ,DATA1_CD_05_exa01celadm01 \       
  ,DATA1_CD_06_exa01celadm01 \       
  ,DATA1_CD_07_exa01celadm01 \       
  ,DATA1_CD_08_exa01celadm01 \       
  ,DATA1_CD_09_exa01celadm01 \       
  ,DATA1_CD_10_exa01celadm01 \       
  ,DATA1_CD_11_exa01celadm01 \       
  size=692288M " 
...

dcli -c exa01celadm14 -l root "cellcli -e alter griddisk  
  DATA1_CD_00_exa01celadm14 \       
  ,DATA1_CD_01_exa01celadm14 \       
  ,DATA1_CD_02_exa01celadm14 \       
  ,DATA1_CD_03_exa01celadm14 \       
  ,DATA1_CD_04_exa01celadm14 \       
  ,DATA1_CD_05_exa01celadm14 \       
  ,DATA1_CD_06_exa01celadm14 \       
  ,DATA1_CD_07_exa01celadm14 \       
  ,DATA1_CD_08_exa01celadm14 \       
  ,DATA1_CD_09_exa01celadm14 \       
  ,DATA1_CD_10_exa01celadm14 \       
  ,DATA1_CD_11_exa01celadm14 \       
  size=692288M " 
```

Chapter 3
Administering Oracle Exadata Storage Server Grid Disks with Oracle ASM

3-28
3. Verify the new size of the grid disks associated with the DATAC1 disk group using the following command:

```
dcli -g cell_group -l root "cellcli -e list griddisk attributes name,size\nwhere name like '%DATAC1.*'"
```

Instead of increasing the size of the DATA disk group, you could instead create new disk groups with the new free space or keep it free for future use. In general, Oracle recommends using the smallest number of disk groups needed (typically DATA, RECO, and DBFS_DG) to give the greatest flexibility and ease of administration. However, there may be cases, perhaps when using virtual machines or consolidating many databases, where additional disk groups or available free space for future use may be desired.

If you decide to leave free space on the grid disks in reserve for future use, please see the My Oracle Support Note 1684112.1 for the steps on how to allocate free space to an existing disk group at a later time.

**Related Topics**

- How to resize ASM disks in Exadata (My Oracle Support Doc ID 1684112.1)
- Determine the Amount of Available Space

To increase the size of the disks in a disk group you must either have unallocated disk space available, or you have to reallocate space currently used by a different disk group.

### 3.3.3.5 Increase the Size of the Oracle ASM Disks

You can increase the size used by the Oracle ASM disks after increasing the space allocated to the associated grid disks.

This task is a continuation of an example where space in the RECO1 disk group is being reallocated to the DATAC1 disk group.

You must have completed the task of resizing the grid disks before you can resize the corresponding Oracle ASM disk group.

1. Increase the Oracle ASM disks for DATAC1 disk group to the new size of the grid disks on the storage cells.

```sql
SQL> ALTER DISKGROUP datac1 RESIZE ALL;
```
This command resizes the Oracle ASM disks to match the size of the grid disks.

**Note:**

If the specified disk group has quorum disks configured within the disk group, then the `ALTER DISKGROUP ... RESIZE ALL` command could fail with error `ORA-15277`. Quorum disks are configured if the requirements specified in *Oracle Exadata Database Machine Maintenance Guide* are met.

As a workaround, you can specify the storage server failure group names (for the ones of FAILURE_TYPE "REGULAR", not "QUORUM") explicitly in the SQL command, for example:

```
SQL> ALTER DISKGROUP datac1 RESIZE DISKS IN FAILGROUP exacell01, exacell02, exacell03;
```

2. Wait for the rebalance operation to finish.

```
SQL> set lines 250 pages 1000
SQL> col error_code form a10
SQL> SELECT dg.name, o.* FROM gv$asm_operation o, v$asm_diskgroup dg
      WHERE o.group_number = dg.group_number;
```

Do not continue to the next step until the query returns zero rows for the disk group that was altered.

3. Verify that the new sizes for the Oracle ASM disks and disk group is at the desired sizes.

```
SQL> SELECT name, total_mb, free_mb, total_mb - free_mb used_mb,
      ROUND(100*free_mb/total_mb,2) pct_free
      FROM v$asm_diskgroup
      ORDER BY 1;
```

```
<table>
<thead>
<tr>
<th>NAME</th>
<th>TOTAL_MB</th>
<th>FREE_MB</th>
<th>USED_MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>116304384</td>
<td>57439796</td>
<td>58864588</td>
</tr>
<tr>
<td>RECOC1</td>
<td>47488896</td>
<td>34542516</td>
<td>12946380</td>
</tr>
</tbody>
</table>
```

```
SQL> SELECT dg.name, d.total_mb, d.os_mb, COUNT(1) num_disks
      FROM v$asm_diskgroup dg, v$asm_disk d
      WHERE dg.group_number = d.group_number
      GROUP BY dg.name, d.total_mb, d.os_mb;
```

```
<table>
<thead>
<tr>
<th>NAME</th>
<th>TOTAL_MB</th>
<th>OS_MB</th>
<th>NUM_DISKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAC1</td>
<td>116304384</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>RECOC1</td>
<td>47488896</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
```
The results of the queries show that the RECOC1 and DATAC1 disk groups and disk have been resized.

Related Topics

- **Determine the Amount of Available Space**
  
  To increase the size of the disks in a disk group you must either have unallocated disk space available, or you have to reallocate space currently used by a different disk group.

### 3.3.4 Determining Which Oracle ASM Disk Group Contains an Oracle Exadata Storage Server Grid Disk

If a grid disk name matches the Oracle ASM disk name, and the name contains the Oracle ASM disk group name, then you can determine the Oracle ASM disk group to which the grid disk belongs.

You can also use SQL commands on the Oracle ASM instance to find the Oracle ASM disk group that matches part of the specific grid disk name. This can help you to determine which Oracle ASM disk group contains a specific grid disk.

**Example 3-6  Determining Grid Disks in an Oracle ASM Disk Group**

This example shows how to find the Oracle ASM disk group that contains grid disks that begin with `DATA0`, for example `DATA0_CD_03_CELL04`.

```sql
SQL> SELECT d.label AS asmdisk, dg.name AS diskgroup
          FROM V$ASM_DISK d, V$ASM_DISKGROUP dg
          WHERE dg.name LIKE 'DATA0%'
              AND d.group_number = dg.group_number;
```

<table>
<thead>
<tr>
<th>ASMDISK</th>
<th>DISKGROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA0_CD_00_CELL04</td>
<td>DATA0</td>
</tr>
<tr>
<td>DATA0_CD_01_CELL04</td>
<td>DATA0</td>
</tr>
<tr>
<td>DATA0_CD_02_CELL04</td>
<td>DATA0</td>
</tr>
<tr>
<td>DATA0_CD_03_CELL04</td>
<td>DATA0</td>
</tr>
</tbody>
</table>

### 3.3.5 Determining Which Oracle Exadata Storage Server Grid Disks Belong to an Oracle ASM Disk Group

If a grid disk name contains the Oracle ASM disk group name, then you can use SQL commands on the Oracle ASM instance to list the Oracle ASM disk group names.

You can use the CellCLI utility to search for specific grid disk names.
Example 3-7 Displaying Oracle ASM Disk Group Names

This example shows how to use a SQL command to display the Oracle ASM disk group names on the Oracle ASM instance.

```
SQL> SELECT name FROM V$ASM_DISKGROUP;
NAME
-------------------------------
CONTROL
DATA0
DATA1
DATA2
LOG
STANDBY
```

Example 3-8 Searching for Grid Disks by Name

This example shows how to display similar grid disk group names on the cell using the dcli utility.

```
$ ./dcli "cellcli -e list griddisk where -c cell04"
data0_CD_01_cell04
data0_CD_02_cell04
data0_CD_03_cell04
...
```

Related Topics

- Naming Conventions for Oracle Exadata Storage Server Grid Disks
  Using a consistent naming convention helps to identify Exadata components.

3.3.6 Handling Disk Replacement

If a disk has a problem, the physical disk status changes.

When a physical disk is removed, its status becomes not present. Oracle ASM may take a grid disk offline when getting I/O errors while trying to access a grid disk on the physical disk. When the physical disk is replaced, Oracle Exadata System Software automatically puts the grid disks on the physical disk online in their respective Oracle ASM disk groups. If a grid disk remains offline longer than the time specified by the disk_repair_time attribute, then Oracle ASM force drops that grid disk and starts a rebalance to restore data redundancy. Oracle ASM monitors the rebalance operation, and Oracle Exadata System Software sends an e-mail message when the operation is complete.

The following table summarizes the physical disk statuses, and how Oracle ASM handles grid disks when the physical disk has a problem.
### Table 3-1 Physical Disk Status

<table>
<thead>
<tr>
<th>Physical Disk Status</th>
<th>Oracle Exadata System Software Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal</td>
<td>Disk is functioning normally</td>
</tr>
<tr>
<td>_not present</td>
<td>Disk has been removed</td>
</tr>
<tr>
<td>predictive failure</td>
<td>Disk is having problems, and may fail</td>
</tr>
<tr>
<td>critical</td>
<td>Disk has failed</td>
</tr>
<tr>
<td>poor performance</td>
<td>Disk is performing poorly</td>
</tr>
</tbody>
</table>

After a physical disk is replaced, Oracle Exadata System Software automatically creates the grid disks on the replacement disk, and adds them to the respective Oracle ASM disk groups. An Oracle ASM rebalance operation relocates data to the newly-added grid disks. Oracle ASM monitors the rebalance operation, and Oracle Exadata System Software sends an e-mail message when the operation is complete.
Related Topics

• Maintaining the Hard Disks of Exadata Storage Servers
4 Administering Oracle Database on Exadata

- Administering SQL Processing Offload
- Overriding the Default Caching Policy
- Administering In-Memory Columnar Caching
- Administering Exadata Hybrid Columnar Compression
- Administering Oracle Database Features on Exadata

4.1 Administering SQL Processing Offload

To optimize the performance of queries that do table and index scans, the database can offload data search and retrieval processing to the Exadata storage servers. This feature is controlled by database initialization parameters:

- **CELL_OFFLOAD_PROCESSING**
  
  The CELL_OFFLOAD_PROCESSING initialization parameter enables SQL processing offload to Oracle Exadata Storage Server.

- **CELL_OFFLOAD_PLAN_DISPLAY**
  
  The database parameter CELL_OFFLOAD_PLAN_DISPLAY determines whether the SQL EXPLAIN PLAN command displays the predicates that can be evaluated by Oracle Exadata System Software as STORAGE predicates for a given SQL command.

- **CELL_OFFLOAD_DECRYPTION**

4.1.1 CELL_OFFLOAD_PROCESSING

The CELL_OFFLOAD_PROCESSING initialization parameter enables SQL processing offload to Oracle Exadata Storage Server.

When the value of the parameter is set to TRUE, predicate evaluation can be offloaded to cells. The default value of the parameter is TRUE. If the parameter is set to FALSE at the session or system level, then the database performs all the predicate evaluation with cells serving blocks. You can set CELL_OFFLOAD_PROCESSING dynamically with the SQL ALTER SYSTEM or ALTER SESSION commands, for example:

```
SQL> ALTER SESSION SET CELL_OFFLOAD_PROCESSING = TRUE;
```

The CELL_OFFLOAD_PROCESSING parameter can also be set with the OPT_PARAM optimizer hint to enable or disable predicate filtering for a specific SQL command.

- To disable CELL_OFFLOAD_PROCESSING for a SQL command:

  ```
  SELECT /*+ OPT_PARAM('cell_offload_processing' 'false') */ COUNT(*) FROM EMPLOYEES;
  ```
To enable CELL_OFFLOAD_PROCESSING for a SQL command:

```sql
SELECT /*+ OPT_PARAM('cell_offload_processing' 'true') */ COUNT(*)
FROM EMPLOYEES;
```

**Note:**

The CELL_OFFLOAD_PROCESSING initialization parameter cannot be used to compare the performance of Oracle Exadata Storage Server with conventional storage. Even when CELL_OFFLOAD_PROCESSING is set to `FALSE`, Oracle Exadata Storage Server has many advantages over conventional storage. Oracle Exadata Storage Server is highly optimized for fast processing of large queries. It has no bottlenecks at the controller or other levels inside the cell. Oracle Exadata System Software uses a modern scale-out architecture and a state-of-the-art InfiniBand network that has much higher throughput than conventional storage networks. Oracle Exadata System Software is tightly integrated with the Oracle Database, and has unique capabilities for setup, execution, monitoring, diagnostics, resource management, and corruption prevention.

**Related Topics**

- *Oracle Database SQL Language Reference*

### 4.1.2 CELL_OFFLOAD_PLAN_DISPLAY

The database parameter `CELL_OFFLOAD_PLAN_DISPLAY` determines whether the SQL EXPLAIN PLAN command displays the predicates that can be evaluated by Oracle Exadata System Software as `STORAGE` predicates for a given SQL command.

The values for the `CELL_OFFLOAD_PLAN_DISPLAY` parameter are `AUTO`, `ALWAYS`, or `NEVER`. The default value is `AUTO`.

- **AUTO** instructs the SQL EXPLAIN PLAN command to display the predicates that can be evaluated as `STORAGE` only if a cell is present and if a table is on the cell.
- **ALWAYS** produces changes to the SQL EXPLAIN PLAN command based on Oracle Exadata System Software, whether or not Oracle Exadata System Software is present or the table is on the cell. You can use this setting to see what can be offloaded to Oracle Exadata Storage Server before migrating to Oracle Exadata Storage Server.
- **NEVER** produces no changes to the SQL EXPLAIN PLAN command for Oracle Exadata System Software.

You can set the `CELL_OFFLOAD_PLAN_DISPLAY` parameter dynamically with the SQL ALTER SYSTEM or ALTER SESSION commands. For example:

```
SQL> ALTER SESSION SET cell_offload_plan_display = ALWAYS;
```

**Related Topics**

- Monitoring Smart I/O Using SQL Explain Plan
4.1.3 CELL_OFFLOAD_DECRYPTION

The CELL_OFFLOAD_DECRYPTION initialization parameter enables decryption offload to Oracle Exadata Storage Servers. This decryption applies to both encrypted tablespaces and encrypted columns. When the value of the parameter is set to TRUE, decryption can be offloaded to cells. The default value of the parameter is TRUE. If the parameter is set to FALSE at the system level, then the database performs all decryption with cells serving blocks. You can set CELL_OFFLOAD_DECRYPTION dynamically with the SQL ALTER SYSTEM command. For example:

```sql
SQL> ALTER SYSTEM SET CELL_OFFLOAD_DECRYPTION = FALSE;
```

4.2 Overriding the Default Caching Policy

Though typically not required, you can use the CELL_FLASH_CACHE segment storage option to override the automatic caching policy for Exadata Smart Flash Cache.

The CELL_FLASH_CACHE option supports the following settings:

- **NONE**: This value ensures that Exadata Smart Flash Cache never caches the corresponding segment. By using this setting on peripheral database segments, more cache space is available for more-important and frequently-accessed database segments.

- **DEFAULT**: This value specifies that database segments are cached using the default LRU (least recently used) algorithm of Exadata Smart Flash Cache. This value is the default setting for CELL_FLASH_CACHE.

- **KEEP**: This value elevates the segment priority in Exadata Smart Flash Cache. By using this setting, you can increase the likelihood of keeping data from the corresponding segment in the cache.

**Example 4-1 Using CREATE TABLE with CELL_FLASH_CACHE**

The storage clause can be provided during the CREATE and ALTER commands for a table or other objects. This example shows the CREATE TABLE command for CELL_FLASH_CACHE use with the storage clause.

```
CREATE TABLE ptable (c1 number, c2 clob) TABLESPACE TBS_1
PARTITION BY RANGE(c1) ( PARTITION p1 VALUES LESS THAN (100)
  TABLESPACE TBS_2 STORAGE (CELL_FLASH_CACHE DEFAULT),
  PARTITION p2 VALUES LESS THAN (200) TABLESPACE TBS_3
  STORAGE (CELL_FLASH_CACHE KEEP));
```

```
CREATE TABLE tkbcsrbc (c1 number, l1 clob) lob (l1) STORE AS securefile
  (cache nologging STORAGE (CELL_FLASH_CACHE NONE))
PCTFREE 0 TABLESPACE tbs_93 STORAGE
  (initial 128K next 128K pctincrease 0);
```
Example 4-2 Using ALTER TABLE with CELL_FLASH_CACHE

For objects where altering the storage clause is allowed, the ALTER command can be used with CELL_FLASH_CACHE, as shown in this example.

```
ALTER TABLE tkbcsrbc STORAGE( CELL_FLASH_CACHE DEFAULT);
```

```
ALTER TABLE tkbcsrbc MODIFY LOB (l1) (STORAGE (CELL_FLASH_CACHE KEEP));
```

Example 4-3 Using Views to Query CELL_FLASH_CACHE Storage Clause

The CELL_FLASH_CACHE storage clause attribute can be queried using database views based on the object involved.

```
SELECT TABLESPACE_NAME, TABLE_NAME, CELL_FLASH_CACHE FROM user_tables WHERE table_name='TKBCSRBC';
```

```
SELECT CELL_FLASH_CACHE FROM ALL_INDEXES WHERE index_name='TKBCIDX';
```

4.3 Administering In-Memory Columnar Caching

The columnar cache is a section of Exadata Smart Flash Cache that stores data in columnar format. When directed by Oracle Database, Exadata automatically uses a portion of the columnar cache to maintain data using the Oracle Database In-Memory format. You do not need to configure anything on Exadata to use this enhancement.

This feature is available if you have licensed the Oracle Database In-Memory option. To enable this feature, use either of the following database instance parameters:

- Set the INMEMORY_SIZE database instance parameter to a value greater than zero.
- Starting with Oracle Database version 19.8.0.0.200714, you can set INMEMORY_FORCE=cellmemory_level.

This option enables you to use In-Memory Columnar Caching in Exadata Smart Flash Cache without a dedicated In-Memory cache in the database instance.

You can use the CELLMEMORY segment option to override the default behavior for In-Memory Columnar Caching in Exadata Smart Flash Cache:

```
SQL> ALTER TABLE table_name [ [ NO ] CELLMEMORY ] [ MEMCOMPRESS FOR [ QUERY | CAPACITY ] [ LOW | HIGH ] ]
```

<table>
<thead>
<tr>
<th>Options and Clauses</th>
<th>Usage Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO CELLMEMORY</td>
<td>Indicates a table is ineligible for the rewrite from 12.1.0.2 columnar flash cache format into the 12.2 Database In-Memory format.</td>
</tr>
<tr>
<td>CELLMEMORY and CELLMEMORY MEMCOMPRESS FOR CAPACITY</td>
<td>Allows a table to be cached in the default Oracle Database 12.2 In-Memory format. You only need to use this clause to undo a previously specified NO CELLMEMORY statement or to change the specified compression level.</td>
</tr>
</tbody>
</table>
Options and Clauses | Usage Description
--- | ---
CELLMEMORY MEMCOMPRESS FOR QUERY | This option indicates that In-Memory column store data should be compressed less than if MEMCOMPRESS FOR CAPACITY was specified. This option can provide an increase in query performance but requires almost twice as much flash space.
LOW and HIGH | Not implemented at this time.

Example 4-4  Using CELLMEMORY and INMEMORY Options on the Same Table

You can use both INMEMORY and CELLMEMORY on the same table. For example:

```sql
CREATE TABLE t (c1 NUMBER) INMEMORY CELLMEMORY MEMCOMPRESS FOR QUERY;
```

Specifying both options is useful for when you have a low priority table that is unlikely to get loaded in memory. By also specifying CELLMEMORY you still get columnar performance.

4.4 Administering Exadata Hybrid Columnar Compression

Use these procedures to administer Oracle Database objects that use Exadata Hybrid Columnar Compression.

- **Determining If a Table Is Compressed**
  Query the `*_TABLES` or `*_TAB_PARTITIONS` data dictionary views to determine whether a table or partitioned is compressed.

- **Determining Which Rows are Compressed**
  When Exadata Hybrid Columnar Compression tables are updated, the rows change to a lower level of compression.

- **Changing Compression Level**
  You can change the compression level for a partition, table, or tablespace.

- **Importing and Exporting Exadata Hybrid Columnar Compression Tables**
  You can use the `impdp` and `expdp` commands to import and export Exadata Hybrid Columnar Compression tables.

- **Restoring an Exadata Hybrid Columnar Compression Table**
  The compressed table backup can be restored to a system that supports Exadata Hybrid Columnar Compression, or to a system that does not support Exadata Hybrid Columnar Compression.

4.4.1 Determining If a Table Is Compressed

Query the `*_TABLES` or `*_TAB_PARTITIONS` data dictionary views to determine whether a table or partitioned is compressed.

- Query the `*_TABLES` data dictionary views to determine table compression.
In the *_TABLES data dictionary views, compressed tables have enabled in the COMPRESSION column.

```
SQL> SELECT table_name, compression, compress_for FROM user_tables;
                   table_name       compression   compress_for
----------------- ------------  -----------------
                     T1            DISABLED
                     T2             ENABLED       BASIC
                     T3             ENABLED       OLTP
                     T4             ENABLED       QUERY HIGH
                     T5             ENABLED       ARCHIVE LOW
```

For partitioned tables, the COMPRESSION column is NULL in the *_TABLES data dictionary views.

- For partitioned tables, query the *_TAB_PARTITIONS data dictionary views.

```
The COMPRESSION column of the *_TAB_PARTITIONS views indicates the table partitions that are compressed. The COMPRESS_FOR column indicates the compression method in use for the table or partition.
```

```
SQL> SELECT table_name, partition_name, compression, compress_for FROM user_tab_partitions;
                   table_name   partition_name   compression   compress_for
----------------- ----------------  ------------  -----------------
                     SALES       Q4_2004          ENABLED       ARCHIVE HIGH
                     SALES       Q3_2008          ENABLED       QUERY HIGH
                     SALES       Q4_2008          ENABLED       QUERY HIGH
                     SALES       Q1_2009          ENABLED       OLTP
                     SALES       Q2_2009          ENABLED       OLTP
```

### 4.4.2 Determining Which Rows are Compressed

When Exadata Hybrid Columnar Compression tables are updated, the rows change to a lower level of compression.

For example, the compression level might change from COMP_FOR_QUERY_HIGH to COMP_FOR OLTP or COMP_NOCOMPRESS.

By sampling the table rows, you can determine the percentage of rows that are no longer at the higher compression level.

- Use the following query to determine the compression level of a row:

```sql
DBMS_COMPRESSION.GET_COMPRESSION_TYPE (ownname IN VARCHAR2, tabname IN VARCHAR2, row_id IN ROWID)
RETURN NUMBER;
```
You can use `ALTER TABLE` or `MOVE PARTITION` to set the rows to a higher compression level. For example, if 10 percent of the rows are no longer at the highest compression level, then you might alter or move the rows to a higher compression level.

**Related Topics**
- GET_COMPRESSION_TYPE Function

### 4.4.3 Changing Compression Level

You can change the compression level for a partition, table, or tablespace.

The following example describes a scenario when you might want to change the compression level.

A company uses warehouse compression for its sales data, but sales data older than six months is rarely accessed. If the sales data is stored in a table that is partitioned based on the age of the data, then the compression level for the older data can be changed to archive compression to free up disk space.

- To change the compression level of a partitioned table you can use the `DBMS_REDEFINITION` package.

  This package performs online redefinition of a table by creating a temporary copy of the table which holds the table data while it is being redefined. The table being redefined remains available for queries and DML statements during the redefinition. The amount of free space for online table redefinition depends on the relative compression level for the existing table, and the new table. Ensure you have enough hard disk space on your system before using the `DBMS_REDEFINITION` package.

- To change the compression level for a single partition of a partitioned table, you can use the `ALTER TABLE ... MODIFY PARTITION` command.

- To change the compression level of a non-partitioned table use the `ALTER TABLE ... MOVE command with the COMPRESS FOR clause`.

  To perform DML statements against the table while the `ALTER TABLE ... MOVE` command is running, you must also add the `ONLINE` clause.

- To change the compression level for a tablespace, use the `ALTER TABLESPACE` command.

  This defines the default for new objects created in the tablespace. Existing objects are not changed or moved.

- You can use Automatic Data Optimization (ADO) to create policies that automatically adjust the compression level.

**Related Topics**
- Oracle Database Administrator’s Guide
- Oracle Database PL/SQL Packages and Types Reference
- Using Automatic Data Optimization
4.4.4 Importing and Exporting Exadata Hybrid Columnar Compression Tables

You can use the **impdp** and **expdp** commands to import and export Exadata Hybrid Columnar Compression tables.

Exadata Hybrid Columnar Compression tables can be imported using the **impdp** command of the Data Pump Import utility. By default, the **impdp** command preserves the table properties and the imported table is Exadata Hybrid Columnar Compression table. The tables can also be exported using the **expdp** command.

On tablespaces not supporting Exadata Hybrid Columnar Compression, the **impdp** command fails with the following error:

*ORA-64307: Exadata Hybrid Columnar Compression is not supported for tablespaces on this storage type*

You can import the Exadata Hybrid Columnar Compression table as an uncompressed table using the **TRANSFORM:SEGMENT_ATTRIBUTES=n** option clause of the **impdp** command.

An uncompressed or OLTP-compressed table can be converted to Exadata Hybrid Columnar Compression format during import. To convert a non-Exadata Hybrid Columnar Compression table to an Exadata Hybrid Columnar Compression table, do the following:

1. Specify default compression for the tablespace using the **ALTER TABLESPACE ... SET DEFAULT COMPRESS** command.
2. Override the **SEGMENT_ATTRIBUTES** option of the imported table during import.

**Related Topics**

- Oracle Data Pump Import
- **ALTER TABLESPACE**

4.4.5 Restoring an Exadata Hybrid Columnar Compression Table

The compressed table backup can be restored to a system that supports Exadata Hybrid Columnar Compression, or to a system that does not support Exadata Hybrid Columnar Compression.

- When restoring a table with Exadata Hybrid Columnar Compression to a system that supports Exadata Hybrid Columnar Compression, restore the file using Oracle Recovery Manager (RMAN).
- When an Exadata Hybrid Columnar Compression table is restored to a system that does not support Exadata Hybrid Columnar Compression, you must convert the table from Exadata Hybrid Columnar Compression to an uncompressed format.

Use the following steps to convert an Exadata Hybrid Columnar Compression table to an uncompressed format.

1. Ensure that there is sufficient space to store the data in uncompressed format.
2. Use RMAN to restore the Exadata Hybrid Columnar Compression tablespace.

3. Alter the table compression from Exadata Hybrid Columnar Compression to NOCOMPRESS:
   For example:

   ```sql
   SQL> ALTER TABLE table_name MOVE ONLINE NOCOMPRESS;
   ```

   Alternatively, you can use the following command to move the data in parallel:

   ```sql
   SQL> ALTER TABLE table_name MOVE ONLINE NOCOMPRESS PARALLEL;
   ```

   If the table partitioned, then alter the compression method for each partition separately. For example:

   ```sql
   SQL> ALTER TABLE table_name MOVE PARTITION partition_name NOCOMPRESS ONLINE;
   ```

   After you convert the table to an uncompressed format, you can optionally use another form of compression, such as OLTP compression or Oracle Database In-Memory compression.

**Related Topics**

- Recovering Tables and Table Partitions from RMAN Backups
- ALTER TABLE

## 4.5 Administering Oracle Database Features on Exadata

Use these guidelines to use and administer the following Oracle Database features in conjunction with Oracle Exadata Database Machine.

- **Using Indexes on Exadata**
  In the past, databases required indexes for good performance. However, Oracle Exadata Database Machine can deliver superior scan rates without using indexes.

- **Using SQL Tuning Advisor on Exadata**
  SQL Tuning Advisor takes one or more SQL statements as input and uses the Automatic Tuning Optimizer to perform SQL tuning on the statements.

- **Using SQL Plan Management on Exadata**
  SQL plan management prevents performance regressions resulting from sudden changes to the execution plan of a SQL statement by recording and evaluating the execution plans of SQL statements over time.

### 4.5.1 Using Indexes on Exadata

In the past, databases required indexes for good performance. However, Oracle Exadata Database Machine can deliver superior scan rates without using indexes.

Review the application execution plans that use indexes to determine if they would run faster with Exadata Smart Scan Offload. To determine if a scan would be faster when there is no index, make the index invisible to the optimizer. An invisible index is maintained by DML operations, but it is not used by the optimizer.
To make an index invisible, use the following command and substitute the name of the index in place of `index_name`:

```
SQL> ALTER INDEX index_name INVISIBLE;
```

4.5.2 Using SQL Tuning Advisor on Exadata

SQL Tuning Advisor takes one or more SQL statements as input and uses the Automatic Tuning Optimizer to perform SQL tuning on the statements.

The output of SQL Tuning Advisor is in the form of advice or recommendations, along with a rationale for each recommendation and its expected benefit. SQL Tuning Advisor provides information about the following:

- Missing and stale statistics
- Better execution plans
- Better access paths and objects
- Better SQL statements

System statistics measure the performance of CPU and storage so that the optimizer can use these inputs when evaluating plans. The database automatically gathers default system statistics, called *noworkload statistics*, at the first instance startup. System statistics ensure the optimizer is aware of Oracle Exadata Database Machine performance.

Use the following SQL command to see if Oracle Exadata-specific statistics have been gathered.

```
SELECT pname, PVAL1 FROM aux_stats$ WHERE pname='MBRC';
```

If the `PVAL1` is null or not set, then the system statistics have not been gathered manually and the default system statistics are being used.

It is not recommended to gather system statistics in Exadata mode for all Exadata environments. Oracle recommends using the defaults for system statistics in most cases. If you have acceptable performance with your current execution plans, then do not gather system statistics in Exadata mode. If the cardinality estimates in your execution plans are accurate, but suffer from the optimizer over-estimating the cost of a full table scan where the full scan performs better, then you should gather system statistics in Exadata mode. For new applications or deployments where the impact can be assessed from the beginning, and dealt with easily if there is a problem, you can gather system statistics in Exadata mode.

The following SQL command gathers system statistics in Exadata mode:

```
exec dbms_stats.gather_system_stats('EXADATA');
```

If your new application or new deployment testing reveals that the default system statistics performed as well or better than using manually collected system statistics, you can reset system statistics to their default values using the `DBMS_STATS.DELETE_SYSTEM_STATS` procedure. After executing this procedure, shut down and then reopen the database.
4.5.3 Using SQL Plan Management on Exadata

SQL plan management prevents performance regressions resulting from sudden changes to the execution plan of a SQL statement by recording and evaluating the execution plans of SQL statements over time.

With SQL Plan management, you build SQL plan baselines composed of a set of existing plans known to be efficient. The SQL plan baselines are then used to preserve performance of corresponding SQL statements, regardless of changes occurring in the system, such as software upgrades or the deployment of new application modules.

You can also use SQL plan management to gracefully adapt to changes such as new optimizer statistics or indexes. You can verify and accept only plan changes that improve performance. SQL plan evolution is the process by which the optimizer verifies new plans and adds them to an existing SQL plan baseline.

Both SQL profiles and SQL plan baselines help improve the performance of SQL statements by ensuring that the optimizer uses only optimal plans. Typically, you create SQL plan baselines before significant performance problems occur. SQL plan baselines prevent the optimizer from using suboptimal plans in the future. The database creates SQL profiles when you invoke SQL Tuning Advisor, which you do typically only after a SQL statement has shown high-load symptoms. SQL profiles are primarily useful by providing the ongoing resolution of optimizer mistakes that have led to suboptimal plans.

The `DBMS_SPM` package supports the SQL plan management feature by providing an interface for you to perform controlled manipulation of plan history and SQL plan baselines maintained for various SQL statements. The `DBMS_SPM` package provides procedures and functions for plan evolution.

Starting with Oracle Exadata System Software release 19.1, there is a new parameter `AUTO_SPM_EVOLVE_TASK` for `DBMS_SPM.CONFIGURE`, which can be used with only Exadata Database Machine on-premises and Oracle Cloud deployments. The `AUTO_SPM_EVOLVE_TASK` parameter can have one of three values:

- **ON**: The feature is enabled. The SPM evolve advisor creates a task to periodically manage the SQL plan history. The task determines whether there are alternatives and if SQL execution plans should be evolved and accepted. The task runs outside the normal maintenance window in a similar manner to high-frequency statistics gathering.
- **OFF**: The feature is disabled. This is the default value.
- **AUTO**: Oracle Database decides when to use the feature.
Related Topics

- Overview of SQL Plan Management
- DBMS_SPM
This section explains how to maintain Oracle Exadata System Software.

Caution:
All operations in this chapter must be performed with extreme caution and only after you have ensured you have complete backups of the data. If not, then you may experience irrecoverable data loss.

- **Recommendations for Changing the Exadata Storage Server Network Address**
  Review the following advice before changing the fundamental configuration of a storage server, such as changing the IP address, host name, or RDMA Network Fabric address.

- **Using the ipconf Utility**
  The `ipconf` utility is used to set and change the following parameters on Oracle Exadata servers.

- **Oracle Exadata System Software Validation Tests and Utilities**
  You can use a variety of commands and utilities to validate the Oracle Exadata System Software and hardware configurations.

- **Locating Serial Numbers for System Components**
  You may need to provide the serial numbers for the system components when contacting Oracle Support Services.

- **Diagnostic and Repair Utilities**
  Oracle Exadata System Software includes utilities for diagnostics and repair of Oracle Exadata Storage Server.

- **System Diagnostics Data Gathering with sosreports and Oracle ExaWatcher**
  You can use the sosreport utility and Oracle ExaWatcher to diagnose problems with your system.

- **Serial Console Support**
  The storage servers and database servers of Oracle Exadata Database Machine are configured to provide serial console access.

- **Oracle Linux Kernel Crash Core Files**
  The storage servers and database servers of Oracle Exadata Database Machine are configured to generate Oracle Linux kernel crash core files in the `/var/crash` directory, when the Oracle Linux operating system malfunctions or crashes.

- **Monitoring syslog Messages Remotely**
  By default, storage server syslog messages are written to local log files.
5.1 Recommendations for Changing the Exadata Storage Server Network Address

Review the following advice before changing the fundamental configuration of a storage server, such as changing the IP address, host name, or RDMA Network Fabric address.

- Before changing the storage server configuration, ensure that all Oracle Automatic Storage Management (Oracle ASM), Oracle Real Application Clusters (Oracle RAC) and database instances that use the storage servers do not access the storage server while you are changing the IP address.

- After changing the storage server configuration, ensure that consumers of storage server services are correctly reconfigured to use the new connect information of the storage server. If Oracle Auto Service Request (ASR) is being used, then deactivate the asset from Oracle ASR Manager, and activate the asset with the new IP address.

- When changing a storage server configuration, change only one storage server at a time to ensure that Oracle ASM and Oracle RAC work properly during the changes.

Related Topics
- About Oracle Auto Service Request

5.2 Using the ipconf Utility

The `ipconf` utility is used to set and change the following parameters on Oracle Exadata servers.

During initial configuration of Oracle Exadata Database Machine, the utility also configures the database servers.

- IP address
- Host name
- NTP server
- Time zone
- DNS name servers
- RDMA Network Fabric addresses

The `ipconf` utility makes a back up copy of the files it modifies. When the utility is rerun, it overwrites the existing backup file. The log file maintains the complete history of every `ipconf` operation performed.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>no option</td>
<td>Utility starts in main editing mode.</td>
</tr>
</tbody>
</table>
### Table 5-1 (Cont.) ipconf Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-check-consistency [-pkey-file pkey.conf]</td>
<td>Determine if pkey is configured on current host by looking into cell.conf. If pkey is configured, this command reversely checks the sub interfaces and the ifcfg config files, compares them with the pkey.conf, and reports any inconsistencies it finds.</td>
</tr>
<tr>
<td>-ignoremismatch</td>
<td>Starts utility in main editing mode when there is a mismatch between the stored cell configuration and the running configuration.</td>
</tr>
<tr>
<td>-ilom print</td>
<td>Prints basic ILOM settings.</td>
</tr>
<tr>
<td>-ilom set</td>
<td>Sets basic ILOM settings.</td>
</tr>
<tr>
<td>-pkey-add $PKEY.CONF -pkey-apply [-force]</td>
<td>Apply the pkey configuration to an unconfigured system or add a new pkey configuration to the current pkey configured host (without changing the current pkey settings). If the interface specified in the pkey file is already, the command exits with an error. If you include the -force option, then the current pkey configuration for the already configured interface is deleted, and the configuration based on the new pkey.conf file is applied.</td>
</tr>
<tr>
<td>-pkey-delete $PKEY.CONF -pkey-apply</td>
<td>To delete some pkeys on the current pkey configured host, where $PKEY.CONF is a simple pkey file that includes only the pkey and physical devices in the entries. If the system is not already configured with pkeys, the command exits with an error.</td>
</tr>
<tr>
<td>-pkey-getruntime</td>
<td>Get the InfiniBand partitioning pkey configuration for the current system.</td>
</tr>
<tr>
<td>-pkey-matchruntime [-pkey-file pkey.conf]</td>
<td>Get the run-time pkey configuration, and match it with the file specified with the -pkey-file option. If a file is not specified, then /opt/oracle.cellos/pkey.conf is used as the default.</td>
</tr>
<tr>
<td>-preconf preconf.scv [-pkey-file pkey.conf] { -generate</td>
<td>-generateall</td>
</tr>
<tr>
<td>-semantic</td>
<td>Checks only the DNS and NTP configuration.</td>
</tr>
<tr>
<td>-semantic-min</td>
<td>Checks for access to at least one NTP and one DNS server.</td>
</tr>
</tbody>
</table>
Table 5-1  (Cont.) ipconf Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-update [-dns dns_ip_list] [-ntp ntp_ip_list] [-ilom-dns ilom_dns_list] [-ilom-ntp ilom_ntp_list] [-force</td>
<td>-dry]</td>
</tr>
<tr>
<td>-verbose</td>
<td>The -verbose option shows all details. If -verbose is not used, then only errors are displayed.</td>
</tr>
<tr>
<td>-verify</td>
<td>Verifies the consistency between the stored cell configuration and the running configuration. Success returns zero errors. You can use the following options with -verify: -pkey-file, -semantic, -semantic-min</td>
</tr>
</tbody>
</table>

The following example shows the display for the ipconf utility when setting the Sun ILOM interface.

```
Example 5-1  Using the ipconf Utility to Set the Sun ILOM Interface

# ipconf
Logging started to /var/log/cellos/ipconf.log
Interface ib0 is Linked. hca: mxx4_0
Interface ib1 is Linked. mxx4_0
Interface eth0 is Linked. driver/mac: igb/00:00:00:01:cd:01
Interface eth1 is ... Unlinked. driver/mac: igb/00:00:00:01:cd:02
Interface eth2 is ... Unlinked. driver/mac: igb/00:00:00:01:cd:03
Interface eth3 is ... Unlinked. driver/mac: igb/00:00:00:01:cd:04

Network interfaces
Name  State      IP address      Netmask         Gateway
Hostname
ib0    Linked
ib1    Linked

eth0   Unlinked
eth1   Unlinked
```
Warning. Some network interface(s) are disconnected. Check cables and switches
and retry
Do you want to retry (y/n) [y]: n

The current nameserver(s): 192.0.2.10 192.0.2.12 192.0.2.13
Do you want to change it (y/n) [n]:
The current timezone: America/Los_Angeles
Do you want to change it (y/n) [n]:
The current NTP server(s): 192.0.2.06 192.0.2.12 192.0.2.13
Do you want to change it (y/n) [n]:

Network interfaces
Name  State   IP address      Netmask       Gateway
Hostname
eth0  Linked  192.0.2.151  255.255.252.0  192.0.2.15  Manager
myg.example.com
eth1  Unlinked
eth2  Unlinked
eth3  Unlinked
bond0 ib0,ib1  192.168.13.101  255.255.252.0  Private
myg.example.com

Select interface name to configure or press Enter to continue:

Select canonical hostname from the list below
1: myg.example.com
2: myg-private.example.com
Canonical fully qualified domain name [1]:

Select default gateway interface from the list below
1: eth01
Default gateway interface [1]:

Canonical hostname: myg.example.com
Nameservers: 192.0.2.10 192.0.2.12 192.0.2.13
Timezone: America/Los_Angeles
NTP servers: 192.0.2.06 192.0.2.12 192.0.2.13

Network interfaces
Name  State  IP address      Netmask        Gateway
Hostname
eth0  Linked 192.0.2.151 255.255.252.0  192.0.2.15  myg.example.com
eth1  Unlinked
eth2  Unlinked
eth3  Unlinked
bond0 ib0,ib1  192.168.13.101 255.255.252.0 Private myg-
priv.example.com

Is this correct (y/n) [y]:

Chapter 5
Using the ipconf Utility
Do you want to configure basic ILOM settings (y/n) [y]: y
Loading configuration settings from ILOM ...
ILOM Fully qualified hostname [myg_ilom.example.com]:
ILOM IP discovery (static/dhcp) [static]:
ILOM IP address [192.0.2.201]:
ILOM Netmask [255.255.252.0]:
ILOM Gateway or none [192.0.2.15]:
ILOM Nameserver or none: [192.0.2.10]:
ILOM Use NTP Servers (enabled/disabled) [enabled]:
ILOM First NTP server. Fully qualified hostname or ip address or none [192.0.2.06]:
ILOM Second NTP server. Fully qualified hostname or ip address or none [none]:

Basic ILOM configuration settings:
Hostname : myg.example.com
IP Discovery : static
IP Address : 192.0.2.10
Netmask : 255.255.252.0
Gateway : 192.0.2.15
DNS servers : 192.0.2.10
Use NTP servers : enabled
First NTP server : 192.0.2.06
Second NTP server : none
Timezone (read-only) : America/Los Angeles

Is this correct (y/n) [y]:

Related Topics

- Set the NTP Server on Exadata Storage Servers
- Set the NTP Server Address on the Database Servers
- Change the DNS Server Address on the Database Server
- Change the DNS Server on Exadata Storage Server

5.3 Oracle Exadata System Software Validation Tests and Utilities

You can use a variety of commands and utilities to validate the Oracle Exadata System Software and hardware configurations.

- **Summary of Software and Firmware Components on Oracle Exadata Storage Servers**
  The imageinfo command located in the /usr/local/bin/ directory provides a summary of release and status of the software and firmware components on Oracle Exadata Storage Servers.

- **Oracle Exadata Storage Server Image History**
  The imagehistory command lists the version history for Oracle Exadata Storage Server.
• **Validation of the State and Health of the System**
  The validation framework runs different tests under certain conditions, such as on first boot after recovery of an Oracle Exadata Storage Server using the rescue and recovery functionality of the CELLBOOT USB flash drive, or when patching an Oracle Exadata Storage Server.

### 5.3.1 Summary of Software and Firmware Components on Oracle Exadata Storage Servers

The `imageinfo` command located in the `/usr/local/bin/` directory provides a summary of release and status of the software and firmware components on Oracle Exadata Storage Servers.

The software and firmware components comprise the storage server image. The release and status information is required when working with My Oracle Support.

The following table lists the output fields from the `imageinfo` command.

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active image activated</td>
<td>Date stamp in UTC format when the image on the cell was considered completed, either successfully or unsuccessfully. A cell patch updates the time stamp to indicate the time the cell was patched.</td>
</tr>
<tr>
<td>Active image status</td>
<td>Status of the cell image based on the success or failure of a set of self-tests and configuration actions, collectively known as <strong>validations</strong>. When this status is undefined, empty or failure, then examine the different validation logs in the <code>/var/log/cellos</code> directory to determine the cause for the status.</td>
</tr>
<tr>
<td>Active image version</td>
<td>Main release version of the overall cell image indicating a specific combination of releases of operating system, core Oracle Exadata System Software (the cell rpm), and the firmware levels for most key components of the cell. A cell patch usually updates this information. The first five separated fields of the version match the standard way Oracle product releases are identified. The last field is the exact build number of the release. It corresponds to <strong>YYMMDD</strong> format of the build date.</td>
</tr>
<tr>
<td>Active system partition on device</td>
<td>Cell operating system root (/) partition device. A typical successful cell patch switches the cell from its active partitions to inactive partitions. Each successful cell patch keeps the cell switching between the active and inactive partitions. There are few occasions when the cell patch does not switch partitions. These are rare, and are known as <strong>in-partition patches</strong>.</td>
</tr>
</tbody>
</table>
Table 5-2  (Cont.) Description of imageinfo Command Output

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boot area has rollback archive for version</td>
<td>For a patched cell using non in-partition cell patch, this indicates whether there is a suitable back up archive that can be used to roll the cell back to the inactive image version. Existence of this archive is necessary but not sufficient for rolling back to inactive version of the cell image.</td>
</tr>
<tr>
<td>Cell boot usb partition</td>
<td>Oracle Exadata Storage Server boot and rescue USB partition.</td>
</tr>
<tr>
<td>Cell boot usb version</td>
<td>Version of the software on the boot USB. On a healthy cell this release must be identical to the value of the Active image version line.</td>
</tr>
<tr>
<td>Cell rpm version</td>
<td>Cell software version or cell rpm version as reported by the CellCLI utility.</td>
</tr>
<tr>
<td>Cell version</td>
<td>Release version as reported by the CellCLI utility.</td>
</tr>
<tr>
<td>In partition rollback</td>
<td>Some cell patches do not switch the partitions. These are in-partition patches. This field indicates whether there is enough information to roll back such patch.</td>
</tr>
<tr>
<td>Inactive image activated</td>
<td>Time stamp for activation of the inactive image. This field is similar to active image activated field.</td>
</tr>
<tr>
<td>Inactive image status</td>
<td>Status of the inactive image. This field is similar to the status of the active image.</td>
</tr>
<tr>
<td>Inactive image version</td>
<td>Version of the cell before the most-recent patch was applied.</td>
</tr>
<tr>
<td>Inactive software partition on device</td>
<td>Oracle Exadata System Software file system partition, /opt/oracle, for the inactive image.</td>
</tr>
<tr>
<td>Inactive system partition on device</td>
<td>The root (/) file system partition for the inactive image.</td>
</tr>
<tr>
<td>Kernel version</td>
<td>Operating system kernel version of the cell.</td>
</tr>
<tr>
<td>Rollback to inactive partition</td>
<td>Summary indicator for a non-in-partition patched cell indicating whether rollback can be run on the cell to take it back to inactive version of the cell image. On a new cell, this field is empty or has the value undefined.</td>
</tr>
</tbody>
</table>

The following is an example of the output from the imageinfo command:

Kernel version: 2.6.18-194.3.1.0.3.el5 #1 SMP Tue Aug 31 22:41:13 EDT 2010 x86_64
Cell version: OSS_MAIN_LINUX.X64_101105
Cell rpm version: cell-11.2.2.1.1_LINUX.X64_101105-1
Active image version: 11.2.2.1.1.101105
Active image activated: 2010-11-06 21:52:08 -0700
Active image status: success
Active system partition on device: /dev/md5
Active software partition on device: /dev/md7

In partition rollback: Impossible

Cell boot usb partition: /dev/sdm1
Cell boot usb version: 11.2.2.1.1.101105

Inactive image version: 11.2.1.3.1
Inactive image activated: 2010-08-28 20:01:30 -0700
Inactive image status: success
Inactive system partition on device: /dev/md6
Inactive software partition on device: /dev/md8

Boot area has rollback archive for the version: 11.2.1.3.1
Rollback to the inactive partitions: Possible

Related Topics

• **Validation of the State and Health of the System**
  The validation framework runs different tests under certain conditions, such as on first boot after recovery of an Oracle Exadata Storage Server using the rescue and recovery functionality of the CELLBOOT USB flash drive, or when patching an Oracle Exadata Storage Server.

5.3.2 Oracle Exadata Storage Server Image History

The `imagehistory` command lists the version history for Oracle Exadata Storage Server.

For example, if a storage server was updated from release 11.2.1.2.6 to release 11.2.1.3.1, and then updated to release 11.2.1.2.3, the `imagehistory` command displays this history. The following is an example of the output:

```
# imagehistory
Version                  : 11.2.1.2.3
Image activation date    : 2012-12-03 06:06:46 -0700
Imaging mode             : fresh
Imaging status           : success

Version                  : 11.2.3.2.0.120713
Image activation date    : 2012-12-12 17:56:31 -0700
Imaging mode             : out of partition upgrade
Imaging status           : success
```

5.3.3 Validation of the State and Health of the System

The validation framework runs different tests under certain conditions, such as on first boot after recovery of an Oracle Exadata Storage Server using the rescue and recovery functionality of the CELLBOOT USB flash drive, or when patching an Oracle Exadata Storage Server.
Validation framework is a set of validation tests that run at boot time at the rc.local level. The logs for the tests are available in the /var/log/cellos/validations directory.

Health check validations are a set of quick health checks on the system on each boot, such as basic health of the disks, and report the status. If a validation fails, then you should examine the log file for the cause as it may indicate potential problem requiring attention.

Automatic patch rollback occurs if one or more validation checks fail after patch application. Refer to the documentation for the specific patch.

Check for any failures reported in the /var/log/cellos/vldrun.first_boot.log file after the first boot configuration. For all subsequent boots, the /var/log/cellos/validations.log file contains information about failed validations. For each failed validation, perform the following procedure:

1. Look for /var/log/cellos/validations/failed_validation_name.SuggestedRemedy file. The file exists only if the validation process has identified some corrective action. Follow the suggestions in the file to correct the cause of the failure.

2. If the SuggestedRemedy file does not exist, then examine the log file for the failed validation in /var/log/cellos/validations to track down the cause, and correct it as needed.

### 5.4 Locating Serial Numbers for System Components

You may need to provide the serial numbers for the system components when contacting Oracle Support Services.

Serial numbers for system components can be determined by using the following procedure:

1. Log in as the root user.

2. Use the CheckHWnFWProfile command to view the serial numbers.

    # /opt/oracle.SupportTools/CheckHWnFWProfile -action list -mode serial_numbers

Each time the system is booted, the serial numbers are written to the /var/log/cellos/validations/SerialNumbers file. This file can be used as a historic record of the serial numbers. The file also contains configuration information for some components.

### 5.5 Diagnostic and Repair Utilities

Oracle Exadata System Software includes utilities for diagnostics and repair of Oracle Exadata Storage Server.

The utilities help diagnose and repair problems that may occur during the normal life cycle of Oracle Exadata Storage Servers. The utilities are in the /opt/oracle.SupportTools directory.
Note:

All utilities must be run as the root user from the /opt/oracle.SupportTools directory.

• **The CheckHWnFWProfile Utility**  
The CheckHWnFWProfile utility checks that the system meets the required hardware and firmware specifications, and reports any mismatches.

• **The Diagnostic ISO File**  
Use the diagnostic ISO file (diagnostics.iso) to diagnose and recover from serious problems.

• **The ibdiagtools Utilities**  
The most useful utilities of the ibdiagtools utilities are the `verify-topology`, `checkbadlinks.pl`, and `infinicheck` utilities.

• **The make_cellboot_usb Utility**  
The `make_cellboot_usb` utility allows you to rebuild a damaged CELLBOOT USB flash drive.

5.5.1 The CheckHWnFWProfile Utility

The CheckHWnFWProfile utility checks that the system meets the required hardware and firmware specifications, and reports any mismatches.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>.CheckHWnFWProfile</code></td>
<td>When run without options, the utility checks the existing hardware and firmware components against the expected values.</td>
</tr>
<tr>
<td><code>.CheckHWnFWProfile -action list</code></td>
<td>View the existing hardware and firmware versions on the system.</td>
</tr>
<tr>
<td><code>.CheckHWnFWProfile -action alter_config -property HWFW_Checker_Updater_Status -value Disabled</code></td>
<td>Disable the CheckHWnFWProfile utility.</td>
</tr>
<tr>
<td><code>.CheckHWnFWProfile -action alter_config -property HWFW_Checker_Updater_Status -value Enabled</code></td>
<td>Enable the CheckHWnFWProfile utility.</td>
</tr>
<tr>
<td><code>.CheckHWnFWProfile -action check -component list_of_components</code></td>
<td>Check specified components against the expected values.</td>
</tr>
<tr>
<td><code>.CheckHWnFWProfile -action list -component list_of_components</code></td>
<td>View the hardware and firmware versions of specified components on the system.</td>
</tr>
</tbody>
</table>
Table 5-3  (Cont.) CheckHWnFWProfile Utility Commands

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
</table>
| ./CheckHWnFWProfile -action list -mode       | List serial numbers. The list includes the following serial numbers:  
| serial_numbers                               | · System  
|                                               | · Disk controller  
|                                               | · Each disk  
|                                               | · RDMA Network Fabric host channel adapter (HCA)  
|                                               | Depending on the system, serial numbers for all the memory (RAM) modules may be included.                                                   |
| ./CheckHWnFWProfile -action list -mode       | View the expected hardware and firmware.                                                                                                      |
| supported_info                               |                                                                                                                                               |
| ./CheckHWnFWProfile -h                       | View help and utility usage.                                                                                                                                 |
| ./CheckHWnFWProfile --help                   |                                                                                                                                               |

5.5.2 The Diagnostic ISO File

Use the diagnostic ISO file (diagnostics.iso) to diagnose and recover from serious problems.

You may need to boot a server using the diagnostic ISO file to diagnose and recover from serious problems. For example, when the system is inaccessible due to system damage or damage to the CELLBOOT USB flash drive.

Use this facility only as directed to perform specific documented maintenance tasks, or under the guidance of Oracle Support Services.

- **Booting a Server using the Diagnostic ISO File**

  Use the following procedure to boot an Oracle Exadata Storage Server or Oracle Exadata Database Server using the diagnostic ISO file (diagnostics.iso).

1. **Download the diagnostic ISO file (diagnostics.iso)** corresponding to your current Oracle Exadata System Software release.

   If required, use the `imageinfo` command on a running server to determine your current Oracle Exadata System Software release.

   Download the Diagnostic ISO file to a client workstation or server that you will use to access the ILOM interface on the server that you want to boot in diagnostics mode. The client machine must have a Web browser and network access to the ILOM interface on the target server.

   If you are using Oracle Exadata System Software Release 18.x, or earlier, you can find the Diagnostic ISO file at `/opt/oracle.SupportTools/diagnostics.iso` on your Oracle Exadata Storage Servers or Oracle Exadata Database Servers. Otherwise, search the My Oracle Support (MOS) patch repository using "exadata diagnostic iso" as the search term.
You can also locate the Diagnostic ISO file in the Supplemental README that is associated with your Oracle Exadata System Software release. The Supplemental README for each Oracle Exadata System Software release is documented in My Oracle Support document 888828.1.

2. Start the ILOM Web client.

   There are two ways to make a physical connect to the server processor from the client machine: Attach a serial cable between a console and the SER MGT port on the server, or attach an ethernet cable between the network switch and the NET MGT port on the server, then connect the workstation to the network. For more details see Connect to Oracle ILOM

   This procedure assumes you are using a network connected workstation.

   Start a Web browser on your client machine and navigate to http://ILOM_SP_IPaddress. In the URL, ILOM_SP_IPaddress is the IP address of the ILOM service processor on the target server.

   The Oracle ILOM Web client login dialog appears.

3. Log in to the Oracle ILOM Web client using the root account and the password.

4. Use the ILOM Web interface to attach the diagnostics.iso file as a virtual CDROM device.

   a. In the ILOM Web interface navigation hierarchy, click Remote Control, and then click Redirection

      Depending on the version of the ILOM Web interface that you are using, Redirection may appear in the left hand navigation pane or in a strip of options running along the top of the page.

   b. In the Redirection pane, click Use Video Redirection, then Launch Remote Console.

      In some versions of the ILOM Web interface, the Use Video Redirection button does not exist. In that case, go directly to Launch Remote Console.

   c. In the Remote Console window, choose the Storage menu option in the KVMS menu.

   d. In the Storage Devices dialog, click Add. And in the resulting dialog, select the diagnostics.iso file.

      The following screen image shows an example of what you may see.
e. Back in the Storage Devices dialog, select the entry associated with the diagnostics.iso file and click Connect.

After the connection is established, the label on the Connect button in the Storage Devices dialog changes to Disconnect.

Keep in mind the navigation path to the Storage Devices dialog so that you can easily disconnect the diagnostics.iso file after you are done with it.

5. Configure the server to boot from the diagnostics.iso file by using the newly configured virtual CDROM device.

a. In the ILOM Web interface navigation hierarchy, click Host Management, and then click Host Control.

Depending on the version of the ILOM Web interface that you are using, Host Control may appear in the left hand navigation pane or in a strip of options running along the top of the page.

b. From the resulting list of values, select CDROM.

c. Click Save.

Now, when the system is booted, the diagnostics.iso file is used as the default boot image.

6. Back in the Remote Console window, restart the system by using the following command:

```
-> Start /SP/console -script
```

The system now reboots using the diagnostics.iso file as the default boot image.
5.5.3 The ibdiagtools Utilities

The most useful utilities of the ibdiagtools utilities are the `verify-topology`, `checkbadlinks.pl`, and `infinicheck` utilities.

The `verify-topology` utility checks the correctness and health of InfiniBand Network Fabric connections. For example, it can determine if both cables from the server go to the same switch in the Oracle Exadata Rack. When both cables go to the same switch, the system loses the ability to fail over to another switch if the first switch fails.

The `checkbadlinks.pl` utility reports the links that are operating at 5 Gbps. This is usually an indication that the cables are loose, and need to be reseated.

The `infinicheck` utility reports the base RDMA Network Fabric performance between servers in Oracle Exadata Database Machine, such as expected minimum throughput between the database server and storage server, between storage servers, and between a database server and another database server. This utility can help to identify issues in the RDMA Network Fabric. Because the utility runs stress tests on the RDMA Network Fabric, Oracle recommends using the utility when the system is idle and with all cell services shut down. This utility works on both the InfiniBand Network Fabric and RoCE Network Fabric.

**See Also:**
- For detailed information about the ibdiagtools utilities, refer to the `README.txt` file in the `/opt/oracle.SupportTools/ibdiagtools/` directory.
- Sample outputs from each utility are included in the `/opt/oracle.SupportTools/ibdiagtools/SampleOutputs.txt` file.

5.5.4 The make_cellboot_usb Utility

The `make_cellboot_usb` utility allows you to rebuild a damaged CELLBOOT USB flash drive.

Do not connect more than one USB flash drive to the system when running this utility. The utility builds on the first discovered USB flash drive on the system.

**Note:**
- This utility can only be used on Oracle Exadata Storage Server.
To see what is done before rebuilding the USB flash drive:

```bash
cd /opt/oracle.SupportTools
./make_cellboot_usb -verbose
```

To rebuild the USB flash drive, run the command with one of the following options: `-execute`, `-force`, or `-rebuild`.

```bash
./make_cellboot_usb -execute
```

Or:

```bash
./make_cellboot_usb -force
```

Or:

```bash
./make_cellboot_usb -rebuild
```

### 5.6 System Diagnostics Data Gathering with sosreports and Oracle ExaWatcher

You can use the sosreport utility and Oracle ExaWatcher to diagnose problems with your system.

Every time a server is started, system-wide configuration information is collected by the sosreport utility, and stored in the `/var/log/cellos/sosreports` directory. You can generate a new sosreport by running the following command as the root user. The script starts collecting the information 30 minutes after entering the command.

```bash
/opt/oracle.cellos/vldrun -script sosreport
```

In addition, the `/opt/oracle.ExaWatcher` directory contains the Oracle ExaWatcher system data gathering and reporting utilities. Gathered data is stored in archive subdirectories. The following table describes the data gathered at different intervals by the utility:

<table>
<thead>
<tr>
<th>Collector Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CellSrvStat</td>
<td>Cell server status.</td>
</tr>
<tr>
<td>DiskInfo</td>
<td>I/O statistics of the disk, such as successfully completed reads, merged reads, time spent reading, and so on.</td>
</tr>
<tr>
<td>FlashSpace</td>
<td>RAW value of the flash card space. Minimum interval limit is 300 seconds.</td>
</tr>
<tr>
<td>IBCardInfo (Currently not available for X8M systems)</td>
<td>RDMA Network Fabric card information, and status of InfiniBand Network Fabric ports. Minimum interval is 300 seconds.</td>
</tr>
</tbody>
</table>
Table 5-4  (Cont.) Oracle ExaWatcher Collector Names and Descriptions

<table>
<thead>
<tr>
<th>Collector Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBprocs</td>
<td>Commands that check the RDMA Network Fabric card status. Minimum interval is 600 seconds.</td>
</tr>
<tr>
<td>Iostat</td>
<td>CPU statistics, and I/O statistics for devices and partitions.</td>
</tr>
<tr>
<td>Lsof</td>
<td>Files opened by current processes. Minimum interval limit is 120 seconds.</td>
</tr>
<tr>
<td>MegaRaidFW</td>
<td>MegaRaid firmware information, such as battery information. Minimum interval is 86400 seconds.</td>
</tr>
<tr>
<td>Meminfo</td>
<td>Memory management by the kernel.</td>
</tr>
<tr>
<td>Mpsstat</td>
<td>Microprocessor statistics.</td>
</tr>
<tr>
<td>Netstat</td>
<td>Current network connection statistics.</td>
</tr>
<tr>
<td>Ps</td>
<td>Active processes statistics.</td>
</tr>
<tr>
<td>RDSinfo</td>
<td>Availability of cell servers. Interval limit is 30 seconds.</td>
</tr>
<tr>
<td>Slabinfo</td>
<td>Caches for frequently-used objects in the kernel.</td>
</tr>
<tr>
<td>Top</td>
<td>Dynamic, real-time view of the system.</td>
</tr>
<tr>
<td>Vmstat</td>
<td>Virtual memory status.</td>
</tr>
</tbody>
</table>

To use Oracle ExaWatcher, do the following:

1. **As the root user, start the Oracle ExaWatcher processes and service.**

   ```
   # systemctl start ExaWatcher
   ```

2. **Run the Oracle ExaWatcher utility at the root user.**

   ```
   /opt/oracle.ExaWatcher/ExaWatcher.sh [options]
   ```

The following options are available for use with the Oracle ExaWatcher utility:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No options specified</td>
<td>The utility runs using the default options.</td>
</tr>
<tr>
<td>Option</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| `-c | --command 'collector_name ;;  
"default_command" ... "'` | To change the core command to be run on the current group. Only the following core commands can be changed:  
CellSrvStat  
Iostat  
Mpstat  
Netstat  
Ps  
Top  
Vmstat  
Example: `--command 'Vmstat;; "vmstat -a"'` |
| `-createconf | "config_file_to_create" | null | The utility parses all command line inputs, validates them, and creates a configuration file. If the file path and name is not specified, then the utility overwrites the default configuration file. |
| `-d | --disable "collector_name"` | The name of the collector to be disabled on the utility.  
Example: `--disable "Vmstat"` |
| `-e | --end "end_time"` | The ending time for the current group. The default value is 10 years from current time.  
Example: `--end "11/06/2013 12:01:00"` |
| `-fromconf | "configuration_file" | null | The configuration file to use with the Oracle ExaWatcher utility. The default configuration files are as follows:  
`/opt/oracle.ExaWatcher/ExaWatcher.conf` for Oracle Linux |
| `-g | --group` | Starts a new group for gathering data. Other options can be specified with the group option.  
Example: `--interval 10` |
| `-h | --help` | Displays help information. |
| `-i | --interval "interval_length"` | The sampling interval for the current group, in seconds. The default value is 5 seconds.  
Certain collection modules cannot be run every second because the modules consume resources.  
Example: `--interval 10` |
| `-l | --spacelimit` | Sets the limit for the amount of storage space used by the utility. The limit is specified in MB.  
On database servers, the default is 6 GB. On storage servers, the default is 600 MB.  
Example: `--spacelimit 900` |
| `--lastconf` | The most-recent configuration file used with the utility.  
Data is not collected when using this option. |
### 5.7 Serial Console Support

The storage servers and database servers of Oracle Exadata Database Machine are configured to provide serial console access.

The serial console is useful when collecting Oracle Linux kernel traces or creating crash dump files to help diagnose severe malfunctions.

To access the serial console, perform the following procedure:

1. Connect to the Integrated Lights Out Manager (ILOM) using SSH and log in as an ILOM administrator.
2. Run the `start /SP/console` command
   
   To stop using the console, use the `stop /SP/console` command.

5.8 Oracle Linux Kernel Crash Core Files

The storage servers and database servers of Oracle Exadata Database Machine are configured to generate Oracle Linux kernel crash core files in the `/var/crash` directory, when the Oracle Linux operating system malfunctions or crashes.

The `crash` utility can be used to analyze the crash files. The crash files are automatically removed by the ExaWatcher utility so that the files do not occupy more than 10 percent of the free disk space on the file system. Older crash files are removed first.

5.9 Monitoring syslog Messages Remotely

By default, storage server syslog messages are written to local log files.

A separate management server, known as a `loghost server`, can receive syslog messages from Oracle Exadata Storage Servers.

- To monitor the syslog messages remotely, configure the syslog service on the loghost server to listen for incoming syslog messages by setting `SYSLOGD_OPTIONS -r` in the loghost server `/etc/sysconfig/syslog` file.

- Configure each Oracle Exadata Storage Server to forward specified syslog messages to the loghost server using the `ALTER CELL` command.

  The configuration on Oracle Exadata Storage Servers is maintained across restarts and updates.

- The `ALTER CELL VALIDATE SYSLOGCONF` command can be used to test that messages are sent by the cells and received by the loghost server.

Related Topics

- `ALTER CELL`
Managing I/O Resources

I/O Resource Management (IORM) is a tool for managing how multiple databases and the workloads within the databases share the I/O resources of Oracle Exadata System Software.

To manage workloads within a database, the Oracle Database Resource Manager has been enhanced to work with IORM to provide database resource management.

- **Understanding I/O Resource Management (IORM)**
  IORM manages the storage server I/O resources on a per-cell basis. Whenever the I/O requests start to saturate a cell's capacity, IORM schedules incoming I/O requests according to the configured resource plans.

- **About Consumer Groups and Resource Plans**
  Oracle Exadata Database Machine provides out-of-the-box consumer groups and resource plans specifically designed for data warehouses that use Oracle Exadata System Software.

- **About CDB Plans and Pluggable Databases**
  The container database (CDB) can have multiple workloads within multiple PDBs competing for resources.

- **Administering IORM**
  You can perform various administrative tasks related to I/O Resource Management (IORM).

**Related Topics**

- Managing Resources with Oracle Database Resource Manager
- Master Note for Oracle Database Resource Manager (My Oracle Support Doc ID 1339769.1)
- Tool for Gathering I/O Resource Manager Metrics: metric_iorm.pl (My Oracle Support Doc ID 1337265.1)

### 6.1 Understanding I/O Resource Management (IORM)

IORM manages the storage server I/O resources on a per-cell basis. Whenever the I/O requests start to saturate a cell's capacity, IORM schedules incoming I/O requests according to the configured resource plans.

- **About I/O Resource Management (IORM) in Exadata Database Machine**
  IORM provides many features for managing resource allocations. Each feature can be used independently or in conjunction with other features.

- **About Database Resource Management Within a Database**
  Oracle Database Resource Manager enables you to manage workloads within a database.

- **About Interdatabase Resource Management**
  Interdatabase resource management enables you to manage resources for multiple databases using the same shared storage.
6.1.1 About I/O Resource Management (IORM) in Exadata Database Machine

IORM provides many features for managing resource allocations. Each feature can be used independently or in conjunction with other features.

Storage is often shared by multiple types of workloads and databases. The correct balance of storage resources across the multiple databases is seldom achieved because real world workloads are very dynamic. Running multiple types of workloads and databases on shared storage can often lead to performance problems. For example, large parallel queries on one production data warehouse can impact the performance of critical queries on another production data warehouse. You can mitigate these problems by over-provisioning the storage system, but this diminishes the cost-savings benefit of shared storage. You can also schedule non-critical tasks at off-peak hours, but this manual process is laborious.

IORM allows workloads and databases to share Oracle Exadata Storage Servers according to user-defined policies. To manage workloads within a database, you can define database resource plans, using Oracle Database Resource Manager which has been enhanced to manage Oracle Exadata Storage Server I/O resources. Define a container database (CDB) resource plan that allows management for the various pluggable databases (PDBs). To manage multiple databases, you can define interdatabase plans.

IORM schedules I/O by immediately issuing some I/O requests and queuing others. The I/O requests that are immediately issued are from workloads that have not exceeded their resource allocation, according to the resource plans. The I/O requests that are queued are from workloads that have exceeded their resource allocation. They are issued when their workload no longer exceeds its resource allocation or when the cell is operating below capacity. When the cell is operating below capacity, IORM does not queue I/O requests, and lets a workload exceed its resource allocation as the system is not maxed out.

For example, if a production database and test database are sharing Oracle Exadata Storage Server resources, you can configure resource plans that give priority to the production database. In this case, whenever the test database load would affect the production database performance, IORM schedules the I/O requests such that the production database I/O performance is not impacted. This means that the test database I/O requests are queued until they can be issued without disturbing the production database I/O performance.

Flash IORM protects the latency of critical OLTP I/O requests in flash cache. When table scans are running on flash concurrently with OLTP I/O requests, the OLTP latency is impacted significantly. Flash IORM queues and throttles the table scan, and other low priority I/O requests. The critical OLTP I/O requests are never queued. When the flash disks are not busy serving critical OLTP I/O requests, the queued I/O requests are issued based on the resource allocations in the interdatabase plan.
• **IORM Plans**
  With IORM, you can create a category plan (catPlan) or an interdatabase plan (dbPlan).

• **Methods of Assigning Resources to Databases**
  You can assign resources in an IORM plan using allocations, shares, or limits.

### 6.1.1.1 IORM Plans

With IORM, you can create a category plan (catPlan) or an interdatabase plan (dbPlan).

Category resource management is an advanced feature. It is useful when Oracle Exadata Storage Server is supporting multiple databases, and you want to allocate resources primarily by the category of the work being done. For example, suppose all databases have three categories of workloads: OLTP, reports, and maintenance. To allocate the I/O resources based on these workload categories you would use category resource management.

Interdatabase resource management enables you to manage resources across multiple databases. Interdatabase resource management is configured by using the CellCLI utility to create interdatabase plans. The interdatabase plans specify a resource allocation for each database. You should use this feature if you have multiple databases using Oracle Exadata Storage Server resources.

If an interdatabase plan has been configured, then each database can have a database plan and a resource allocation. A database resource plan specifies how the database resource allocation should be divided across its workloads, or consumer groups. If a database does not have a database resource plan enabled, then the database resource allocation is not divided, and all I/O requests from the database are treated as a single workload.

### 6.1.1.2 Methods of Assigning Resources to Databases

You can assign resources in an IORM plan using allocations, shares, or limits.

With the catPlan, you can use **allocations** to specify percentages for I/O distribution for each database or category. With allocations, IORM can manage up to 32 categories or databases. You can allocate resources across 8 different **levels**. For example, you can put batch operations on a different level from sales operations.

For dbPlans, you can also use allocations, or you can instead use either shares or limits to manage resources.

- Use the **share** method to specify the relative priority for a database. A higher share value implies higher priority and more guaranteed I/O resources. This is the recommendation method for dbPlans.

- Use the **limit** method to restrict the I/O utilization for a database to a specified utilization limit. This directive ensures that the database never utilizes I/O resources beyond the specified limits. For example, if a production and test database are sharing Oracle Exadata Storage Server resources, then set a maximum utilization limit for the test database to limit the I/O utilization for that database.

If a maximum utilization limit is specified, then excess capacity is never used by the databases. As a result, it is possible that the disks are running below full capacity when maximum utilization limits are specified.
Use a value from 1 to 100 for the limit to specify the maximum disk I/O utilization limit for databases. This configuration is ideal for pay-for-performance use cases and should not be used to implement fairness.

I/Os issued to hard disks are not subjected to the maximum utilization limit, but I/Os on flash devices are. This is because almost all of the I/Os are served by Exadata Smart Flash Cache. By enforcing the maximum utilization limit only on the flash devices, the system avoids high latencies and outliers due to low utilization limits on the I/Os that miss the flash cache, such as cache misses or population I/Os.

6.1.2 About Database Resource Management Within a Database

Oracle Database Resource Manager enables you to manage workloads within a database.

A database often has many types of workloads. These workloads may differ in their performance requirements and the amount of I/O that they issue. Database resource management is configured at the database level, using Oracle Database Resource Manager to create database resource plans. You should use this feature if you have multiple types of workloads within a database. You can define a policy for specifying how these workloads share the database resource allocations. If only one database is using the Oracle Exadata Storage Server resources, then this is the only resource management feature that you need.

For each database, you can use Oracle Database Resource Manager for the following tasks:

- **Create resource consumer groups**
  Resource consumer groups provide a way to group sessions that comprise a particular workload. For example, if your database is running four different applications, then you can create four consumer groups, one for each application. If your data warehouse has three types of workloads, such as critical queries, normal queries, and ETL (extraction, transformation, and loading), then you can create a consumer group for each type of workload.

- **Map user sessions to consumer groups**
  Once you have created the consumer groups, you must create rules that specify how sessions are mapped to consumer groups. Oracle Database Resource Manager allows you to create mapping rules based on session attributes such as the Oracle user name, the service that the session used to connect to the database, client machine, client program name, client user name, and so on. If you are creating consumer groups for each application and each application has a dedicated service, then you should create mapping rules based on service names. If you want to dedicate a consumer group to a particular set of users, then you should create mapping rules based on their user names. Sessions that are not explicitly assigned to a consumer group are placed in the **OTHER_GROUPS** consumer group.

- **Create CDB resource plans**
  A container database (CDB) resource plan specifies how CPU and I/O resources are allocated among the different pluggable databases (PDBs) that are part of the same container. The CDB plan is created using Oracle Database Resource Manager. The CDB plan contains a directive for each PDB. The directive defines the number of shares that are allocated to that PDB. The shares define the relative priority of that PDB as compared to other PDBs in the plan.
A maximum utilization limit can be specified for a PDB.

A CDB resource plan also lets you specify `memory_min` and `memory_limit` for each PDB. These parameters specify the flash cache minimum and maximum quotas for each PDB.

- **Create resource plans**

  The database resource plan, also known as an **intradatabase resource plan**, specifies how CPU and I/O resources are allocated among consumer groups in its database. The resource plan is created using Oracle Database Resource Manager. It contains a resource allocation directive for each consumer group, which consists of a percentage and a level. You can specify up to eight levels.

  - Consumer groups at level 2 get resources that were not allocated at level 1 or were not consumed by a consumer group at level 1.
  - Consumer groups at level 3 are allocated resources only when some allocation remains from levels 1 and 2.
  - The same rules apply to levels 4 through 8.

  Multiple levels not only provide a way of prioritizing, they also provide a way of explicitly specifying how all primary and leftover resources are to be used. You can construct resource plans that allocate resources across consumer groups using percentages, priorities, or a combination of the two.

  You can also specify a maximum utilization limit for a consumer group. This works in the same way as a maximum utilization limit for a database, and limits the I/O utilization for the consumer group to the specified value.

  In addition to a CDB plan, each PDB can also create a resource plan to manage the workloads running within the PDB. PDBs only support single level plans with a maximum of 8 consumer groups.

- **Enable a resource plan**

  A database resource plan can be manually enabled with the `RESOURCE_MANAGER_PLAN` initialization parameter or automatically enabled with the job scheduler window.

  When you set a database resource plan on the database, a description of the plan is automatically sent to each cell. For Oracle Real Application Clusters (Oracle RAC) database running on Oracle Exadata Database Machine, all instances in the Oracle RAC cluster must be set to the same resource plan. When a new cell is added or an existing cell is restarted, the current plan of the database is automatically sent to the cell. The resource plan is used to manage resources on both the database server and storage servers (cells).

  Background I/Os are scheduled based on their priority relative to the user I/Os. For example, redo writes, and control file reads and writes are critical to performance and are always prioritized above all user I/Os. The database writer process (DBWR) writes are scheduled at the same priority level as the user I/Os. If a resource plan is not enabled for a database, then all user I/Os are treated equally, and background I/Os are treated as described in this paragraph.

  Oracle provides several predefined plans. The most commonly used are `mixed_workload_plan`, `dss_plan`, and `default_maintenance_plan`.

**Related Topics**

- Managing Resources with Oracle Database Resource Manager
6.1.3 About Interdatabase Resource Management

Interdatabase resource management enables you to manage resources for multiple databases using the same shared storage.

An interdatabase plan specifies how resources are allocated by percentage or share among multiple databases for each storage server. The directives in an interdatabase plan specify allocations to databases, rather than consumer groups. The interdatabase plan is configured and enabled with the CellCLI utility at each storage server.

- **About Interdatabase IORM Plan Directives**
  Interdatabase plan directives are specified using the `db_unique_name` of the database as the identifier.

- **Using Interdatabase Plans for Consolidation and Isolation**
  Interdatabase resource management plans help manage resources when consolidating or isolating databases.

- **About Flash Cache Management in Interdatabase Plans**
  I/O Resource Management can provide predictable performance by guaranteeing space in Exadata Smart Flash Cache.

- **About PMEM Cache Management in Interdatabase Plans**
  I/O Resource Management can provide predictable performance by guaranteeing space in the persistent memory (PMEM) cache.

- **Using IORM to Control Database Access to Flash and PMEM Resources**
  You can use IORM to manage access to valuable flash memory and persistent memory (PMEM) resources.

- **About PMEM Cache Management in Interdatabase Plans**
  You can use IORM to manage access to valuable flash memory and persistent memory (PMEM) resources.

- **Tips for Managing Resource Plans**
  Note the following information when creating and managing resource plans.

**Related Topics**

- **Administering Interdatabase Resource Management**
  You can configure an interdatabase or category plan for Oracle Exadata Database Machine using the CellCLI `ALTER IORMPLAN` command.

**6.1.3.1 About Interdatabase IORM Plan Directives**

Interdatabase plan directives are specified using the `db_unique_name` of the database as the identifier.

For example, you might specify that 70 percent of the I/O resources are allocated to database `SALES`, 30 percent to `HR`, and any of their unused allocations are reallocated to the `TEST_SALES` database. The interdatabase plan is similar to a database resource plan, in that each directive consists of an allocation amount and a level from 1 to 8.

For a given plan, the total allocations at any level must be less than or equal to 100 percent. An interdatabase plan differs from a database resource plan in that it cannot contain subplans and only contains I/O resource directives. Only one interdatabase plan can be active on a cell at any given time.

Share-based plans use a relative share instead of percentage allocations and levels. These plans are simpler to implement, but as effective as percentage allocations. Each database is given a share value which is an integer between 1 and 32.
sum of the shares can be greater than 100. Share-based plans support up to 1024 directives within the interdatabase plan. For example, if a critical database, FINANCE, has 4 shares, and a low-priority database, REPORTING, has 1 share, then the FINANCE database is four times more likely to issue I/Os compared to the REPORTING database.

Oracle Exadata System Software uses the IORM and database resource plans together to allocate I/O resources.

- First, the interdatabase plan allocates the I/O resources to individual databases. Any unused resources are reallocated to other databases, as specified by the plan. This is similar to database resource plans.
- Next, the database resource plan for each database allocates the I/O resources to consumer groups. If a database does not have an active database resource plan, all user I/Os are treated the same. Background I/Os are automatically prioritized relative to the user I/Os based on their importance.

As a best practice, you should create a directive for each database that is using the same cell storage. This is done automatically for shared-based plans, but not for percentage allocation plans. Each database that is not explicitly mapped in a share-based plan gets the default share of 1. To ensure that any database without an explicit directive can be managed with percentage allocation plans, create an allocation named OTHER. Databases without explicit directives are managed using the allocation of the OTHER group directive.

Share-based plans use the DEFAULT directive to specify the default values for the different IORM attributes for databases not explicitly named in the plan. The default share is 1, and the limit is set to 100%. Exadata Smart Flash Cache and Exadata Smart Flash Logging are enabled by default and there are no minimum or maximum flash cache quotas setup by default.

If you have databases with the same DB_UNIQUE_NAME in different Oracle ASM clusters, then starting with Oracle Exadata System Software release 19.1.0, you can use the asmcluster attribute to uniquely identify each database in the interdatabase plan.

**Note:**

For Oracle Data Guard deployments only, I/O Resource Management (IORM) does not support specifying directives using db_name unless the role attribute is also specified in the directive.

### 6.1.3.2 Using Interdatabase Plans for Consolidation and Isolation

Interdatabase resource management plans help manage resources when consolidating or isolating databases.

Consider the consolidation of four different applications on the same Oracle Exadata Storage Servers. If all the applications have similar priority requirements, you can use IORM to equally allocate 25% of the I/O resources. If one application is more important than the others, then you can use IORM to give it a larger percentage of the I/O resources.

Maximum utilization limit is also useful in such consolidation scenarios. You can isolate each application in case one application sees a sudden increase in its workload. Use the maximum utilization limit to isolate the applications, such as specifying a maximum
utilization limit of 40% for each application. In such a scenario, each application can utilize a maximum of 40% of the I/O resources, and does not completely monopolize the system.

Another situation where maximum utilization limits make sense is in ensuring a pay-for-performance model. Service providers want to guarantee performance to their customers corresponding to the level of service being purchased. A customer that is not purchasing the gold tier service should not receive gold level performance. You can use IORM resource plans to distribute the I/O resources among customer applications based on their performance model. Flash cache minimum and maximum quotas are also useful in this model.

6.1.3.3 About Flash Cache Management in Interdatabase Plans

I/O Resource Management can provide predictable performance by guaranteeing space in Exadata Smart Flash Cache.

With multiple databases, and pluggable databases (PDBs) sharing the storage, flash cache space becomes a critical resource that needs to be managed. IORM can reserve space for critical databases or PDBs while preventing less important or rouge databases from consuming the entire flash cache.

You can use attributes in the dbPlan to set hard or soft limits for flash cache resources. A hard limit means the database cannot exceed its quota even when the memory cache is not full. A soft limit means that the specified limit can be exceeded if there are available resources.

- **flashCacheMin** — Specifies the minimum amount of flash cache space that is guaranteed for a database, even if the blocks are cold. This is a hard limit.
  
  Because `flashCacheMin` is a guaranteed reservation, the sum of `flashCacheMin` across all directives should be less than the size of the flash cache to ensure that each database gets its respective quota.

- **flashCacheLimit** — Specifies the soft maximum amount of flash cache space that is available to a database. If the flash cache is not full, a database can exceed the `flashCacheLimit` value.

- **flashCacheSize** — Specifies the hard maximum amount of flash cache space that is available to a database. A database is not allowed to exceed the `flashCacheSize` value at any time.

  However, if you set `flashCacheSize` to a value that is lower than the current space occupied by the database, then starting with Oracle Exadata System Software release 20.1.0 excess data is proactively cleared from the cache. Previously, excess data was only removed when overwritten by other data.

  Starting with Oracle Exadata System Software release 19.2.0, `flashCacheSize` is not a guaranteed reservation if the sum of the `flashCacheSize` across all directives is more than the size of the flash cache. In this case, you can also specify `flashCacheMin` to define a guaranteed minimum quota.

Within each database, you can manage the minimum and maximum quotas for PDBs with Oracle Database Resource Manager, in a container database (CDB) resource plan using the `memory_min` and `memory_limit` directives.
6.1.3.4 About PMEM Cache Management in Interdatabase Plans

I/O Resource Management can provide predictable performance by guaranteeing space in the persistent memory (PMEM) cache.

With multiple databases, and pluggable databases (PDBs) sharing the storage, PMEM cache space becomes a critical resource that needs to be managed. IORM can reserve space for critical databases or PDBs while preventing less important or rouge databases from consuming the entire PMEM cache, available in Oracle Exadata Storage Server X8M and later.

You can use attributes in the dbPlan to set hard or soft limits for PMEM cache resources. A hard limit means the database cannot exceed its quota even when the memory cache is not full. A soft limit means that the specified limit can be exceeded if there are available resources.

- `pmemCacheMin` — Specifies the minimum amount of PMEM cache space that is guaranteed for a database, even if the blocks are cold. This is a hard limit.
  
  Because `pmemCacheMin` is a guaranteed reservation, the sum of `pmemCacheMin` across all directives should be less than the size of the PMEM cache to ensure that each database gets its respective quota.

- `pmemCacheLimit` — Specifies the soft maximum amount of PMEM cache space that is available to a database. If the PMEM cache is not full, a database can exceed the `pmemCacheLimit` value.

- `pmemCacheSize` — Specifies the hard maximum amount of PMEM cache space that is available to a database. A database is not allowed to exceed the `pmemCacheSize` value at any time.

  However, if you set `pmemCacheSize` to a value that is lower than the current space occupied by the database, then starting with Oracle Exadata System Software release 20.1.0 excess data is proactively cleared from the cache. Previously, excess data was only removed when overwritten by other data.

  `pmemCacheSize` is not a guaranteed reservation if the sum of the `pmemCacheSize` across all directives is more than the size of the PMEM cache. In this case, you can also specify `pmemCacheMin` to define a guaranteed minimum quota.

Within each database, you can manage the minimum and maximum quotas for PDBs with Oracle Database Resource Manager, in a container database (CDB) resource plan using the `memory_min` and `memory_limit` directives.

6.1.3.5 Using IORM to Control Database Access to Flash and PMEM Resources

You can use IORM to manage access to valuable flash memory and persistent memory (PMEM) resources.

You can use the `flashcache` and `flashlog` directives in the interdatabase IORM plan to control access to Exadata Smart Flash Cache and flash log. You can set `flashcache=off` to prevent the specified database from using the flash cache. Likewise, setting `flashlog=off` prevents the specified database from using the flash log. By default, all databases can use flash memory for caching and logging.
Likewise, you can use the `pmemcache` and `pmemlog` directives in the interdatabase IORM plan to control access to PMEM cache and PMEM log.

These directives allow valuable flash cache, PMEM cache, flash log, and PMEM log resources to be reserved for mission-critical databases, especially in consolidated environments.

### 6.1.3.6 Tips for Managing Resource Plans

Note the following information when creating and managing resource plans.

- If Oracle Exadata Database Machine is only hosting one database, then an interdatabase plan is not needed.
- If an interdatabase plan is not specified, then all databases receive an equal allocation.
- If only one database is mapped to the `OTHER` directive and all other databases have an explicit directive, then Oracle Exadata System Software uses the database resource plan of that database to determine how the allocation of the `OTHER` database is redistributed among the consumer groups in that database.
- If multiple databases are mapped to the `OTHER` directive, then Oracle Exadata System Software does not use Oracle Database Resource Manager for these databases. All of the I/O requests are treated the same.
- For share-based plans, each database gets its own directive even when it is not explicitly named in the plan. Oracle Exadata System Software uses the database resource plan of the database to determine how the allocation is to be distributed between the consumer groups in the database.
- If you have databases with the same `DB_UNIQUE_NAME` but associated with different Oracle ASM clusters, then, starting with Oracle Exadata System Software release 19.1.0, you can use the `asmcluster` attribute to uniquely identify each database when specifying directives.
- If the container database (CDB) plan specifies `memory_min` or `memory_limit`, and the interdatabase plan specifies only `flashcachesize`, then `memory_min` in the CDB plan is ignored.
- If the container database (CDB) plan specifies `memory_min` or `memory_limit`, and the interdatabase plan specifies only `pmemcachesize`, then `memory_min` in the CDB plan is ignored.

### 6.1.4 About I/O Resource Management Profiles

I/O Resource Management (IORM) interdatabase plans support profiles to ease management, and configuration of interdatabase plans for hundreds of databases.

Profiles introduce a way to allocate I/O resources for groups of databases. Profiles are specified as directives for the interdatabase plan, and are configured using the CellCLI utility. A profile directive consists of an identifier (name), and a set of attributes. To differentiate between a database directive and a profile directive, a qualifier attribute called `type` is used. The `type` attribute can be set to either `database` or `profile`. The following is an example of the `type` attribute syntax:

```
CellCLI> ALTER IORMPLAN DBPLAN=((name=gold, share=10, limit=100, type=profile),
```

Chapter 6

Understanding I/O Resource Management (IORM)

6-10
The preceding example contains three directives for profiles, **GOLD**, **SILVER** and **BRONZE**. All databases with `db_performance_profile` set to **GOLD** would automatically get 10 shares, and 100% limit on the cells. Likewise, in the example above, databases with the **SILVER** profile would get 5 shares and 60% limit, and databases with the **BRONZE** profile would get 1 share and 20% limit.

After your create your profiles, you then map new and existing databases to one of the profiles defined in the interdatabase plan. This is done by setting the `db_performance_profile` initialization parameter for each database to the name of the desired profile. You must then restart the database. As with Oracle Database Resource Manager plans, the IORM profile information is automatically pushed to all the storage servers (cells). The following SQL command displays how the initialization parameter can be set for a database:

```
SQL> ALTER SYSTEM SET db_performance_profile=gold SCOPE=spfile;
```

When adding a new database, you can set the `db_performance_profile` parameter and restart the database. The database automatically inherits the profile attributes without having to modify the interdatabase plan. You can also create interdatabase plans with a mix of profile directives, and database directives.

You can view existing profiles using the `LIST IORMPROFILE` command.

Note the following when managing interdatabase profile plans:

- The `db_performance_profile` parameter is not a dynamic parameter, so profile updates require a database restart.
- If the `type` attribute is not specified, then the directive defaults to the `database` directive.
- An interdatabase plan can specify only 8 profile directives, and 1024 database directives.
- Level, allocation, and role cannot be specified with a profile directive.
- The words `OTHER`, and `DEFAULT` are reserved words. A profile name cannot be `OTHER` or `DEFAULT`.
- The `type` attribute cannot be specified with category plans.
- Profiles cannot be specified in conjunction with category plans.
- If multiple databases are mapped to the `OTHER` directive, then Oracle Exadata Storage Server does not use Oracle Database Resource Manager for these databases. All of the I/O requests are treated the same.

**Related Topics**

- `ALTER IORMPLAN`
6.1.5 About Category Resource Management

Categories represent collections of consumer groups across all databases.

Oracle Database Resource Manager enables you to specify a category for every consumer group. You can manage I/O resources based on categories by creating a category plan. For example, you can specify precedence to consumer groups in the interactive category over consumer groups in the batch category for all databases sharing Oracle Exadata Storage Server.

You can add any number of categories, or modify the predefined categories. You should map consumer groups to the appropriate category for all databases that use the same cell storage. Any consumer group without an explicitly specified category defaults to the OTHER category.

When a category plan is enabled, the category plan is first used to allocate resources among the categories. For each category that is selected, the interdatabase plan is used to select those databases that have consumer groups with the selected category. Finally, the database resource plan of the selected database is used to select one of its consumer groups.

Category plans are configured and enabled using the CellCLI utility on the cell. Only one category plan can be enabled at a time. The predefined categories provided in Oracle Database are described in the following table, along with sample percentages.

<table>
<thead>
<tr>
<th>Category Name</th>
<th>Category Description</th>
<th>Level 1 (%)</th>
<th>Level 2 (%)</th>
<th>Level 3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADMINISTRATIVE</td>
<td>For extremely high-priority work, such as urgent administrative tasks. This category is required.</td>
<td>80</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>INTERACTIVE</td>
<td>For high-priority, performance-sensitive work, such as OLTP transactions.</td>
<td>not set</td>
<td>70</td>
<td>not set</td>
</tr>
<tr>
<td>BATCH</td>
<td>For low-priority work, such as noncritical reports and backup.</td>
<td>not set</td>
<td>not set</td>
<td>70</td>
</tr>
<tr>
<td>MAINTENANCE</td>
<td>For low-priority work, such as automated tasks.</td>
<td>not set</td>
<td>not set</td>
<td>10</td>
</tr>
<tr>
<td>OTHER</td>
<td>For all consumer groups that do not have a category label or reference a category that is not in the current category plan. This category is required.</td>
<td>not set</td>
<td>not set</td>
<td>20</td>
</tr>
</tbody>
</table>

The sample plan shown in the above table prioritizes administrative activity across all databases. It also prioritizes interactive activity over batch, maintenance, and other activities. In the sample plan, the following are the resource allocations:

- Level 1 is given 80 percent of the I/O resources. The ADMINISTRATIVE category is the only category in level 1.
- Level 2 is given all resources that were unallocated or unused by level 1. In this example, level 2 is given 20 percent of the I/O resources and any resources
unused by the ADMINISTRATIVE category. The INTERACTIVE category gets 70 percent of the level 2 amount.

- Level 3 categories are given the remaining resources, including those not used by the INTERACTIVE category. Of the remaining resources, the BATCH category gets 70 percent, the OTHER category gets 20 percent, and the MAINTENANCE category gets 10 percent.

All administrative consumer groups in all databases should be mapped to the ADMINISTRATIVE category. All high-priority user activity, such as consumer groups for important OLTP transactions and time-critical reports, should be mapped to the INTERACTIVE category. All low-priority user activity, such as reports, maintenance, and low-priority OLTP transactions, should be mapped to the BATCH, MAINTENANCE, and OTHER categories.

Related Topics

- Administering Database Resource Management
  To set up database resource management, you must use Oracle Database Resource Manager to configure the consumer groups, assign sessions to consumer groups, create a database resource plan, and enable the plan.

6.2 About Consumer Groups and Resource Plans

Oracle Exadata Database Machine provides out-of-the-box consumer groups and resource plans specifically designed for data warehouses that use Oracle Exadata System Software.

These resource plans can be modified to suit the needs of your environment.

The following consumer groups are for data warehouses:

- **ETL_GROUP**: Consumer group for ETL (extract, transform, and load) jobs.
- **DSS_GROUP**: Consumer group for non-critical decision support system (DSS) queries.
- **DSS_CRITICAL_GROUP**: Consumer group for critical DSS queries.

The following resource plans are for data warehouses:

- **DSS_PLAN Resource Plan**
  The DSS_PLAN resource plan is designed for data warehouses that prioritize critical DSS queries over non-critical DSS queries and ETL jobs.

- **ETL_CRITICAL_PLAN Resource Plan**
  The ETL_CRITICAL_PLAN prioritizes ETL over DSS queries.

6.2.1 DSS_PLAN Resource Plan

The DSS_PLAN resource plan is designed for data warehouses that prioritize critical DSS queries over non-critical DSS queries and ETL jobs.

In this plan, SYS_GROUP has the highest priority, followed by DSS_CRITICAL_GROUP, DSS_GROUP, and then a combination of ETL_GROUP and BATCH_GROUP. No consumer group is allowed to consume all the bandwidth.
Table 6-2   DSS_PLAN Resource Plan for Data Warehouses

<table>
<thead>
<tr>
<th>Consumer Group</th>
<th>Level 1 (%)</th>
<th>Level 2 (%)</th>
<th>Level 3 (%)</th>
<th>Level 4 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYS_GROUP</td>
<td>75</td>
<td>not set</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>DSS_CRITICAL_GROUP</td>
<td>not set</td>
<td>75</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>DSS_GROUP</td>
<td>not set</td>
<td>not set</td>
<td>75</td>
<td>not set</td>
</tr>
<tr>
<td>ETL_GROUP</td>
<td>not set</td>
<td>not set</td>
<td>not set</td>
<td>45</td>
</tr>
<tr>
<td>BATCH_GROUP</td>
<td>not set</td>
<td>not set</td>
<td>not set</td>
<td>45</td>
</tr>
<tr>
<td>ORA$DIAGNOSTICS</td>
<td>not set</td>
<td>5</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>ORA$AUTOTASK_SUB_PLAN</td>
<td>not set</td>
<td>5</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>OTHER_GROUPS</td>
<td>not set</td>
<td>not set</td>
<td>not set</td>
<td>10</td>
</tr>
</tbody>
</table>

As shown in the previous table, the DSS_CRITICAL_GROUP group is only allocated 75 percent at level 2. Any unused allocation goes to the next level, not to other consumer groups at the same level. That means that if the DSS_CRITICAL_GROUP group does not completely consume its allocation, then the allocation is not given to the ORA$DIAGNOSTICS or ORA$AUTOTASK_SUBPLAN groups at the same level, but instead the allocation is given to the DSS_GROUP group at level 3.

6.2.2 ETL_CRITICAL_PLAN Resource Plan

The ETL_CRITICAL_PLAN prioritizes ETL over DSS queries. In this plan, the SYS_GROUP group is given 75 percent of the bandwidth. The remaining bandwidth is divided between the other consumer groups in the ratios specified by the level 2 allocations. The ETL_GROUP and DSS_CRITICAL_GROUP groups have a higher allocation (35 percent) than the DSS_GROUP and BATCH_GROUP groups (10 percent).

Table 6-3   ETL_CRITICAL_PLAN Resource Plan for Data Warehouses

<table>
<thead>
<tr>
<th>Consumer Group</th>
<th>Level 1 (%)</th>
<th>Level 2 (%)</th>
<th>Level 3 (%)</th>
<th>Level 4 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYS_GROUP</td>
<td>75</td>
<td>not set</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>DSS_CRITICAL_GROUP</td>
<td>not set</td>
<td>35</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>DSS_GROUP</td>
<td>not set</td>
<td>10</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>ETL_GROUP</td>
<td>not set</td>
<td>35</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>BATCH_GROUP</td>
<td>not set</td>
<td>10</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>ORA$DIAGNOSTICS</td>
<td>not set</td>
<td>3</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>ORA$AUTOTASK_SUB_PLAN</td>
<td>not set</td>
<td>3</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>OTHER_GROUPS</td>
<td>not set</td>
<td>3</td>
<td>not set</td>
<td>not set</td>
</tr>
</tbody>
</table>

6.3 About CDB Plans and Pluggable Databases

The container database (CDB) can have multiple workloads within multiple PDBs competing for resources.
The Oracle Multitenant container database (CDB) supports many user-defined pluggable databases (PDBs). In a CDB, resources are managed at the following levels:

- **CDB level**: Oracle Database Resource Manager manages the workloads for multiple PDBs that are contending for system and CDB resources. The administrator can specify how resources are allocated to PDBs, and can limit the resource utilization of specific PDBs.

- **PDB level**: Oracle Database Resource Manager manages the workloads within each PDB.

The following CDB plan contains three PDBs named SALES, SERVICES and HR. The PDBs have different shares and max utilization limits in the CDB plan.

<table>
<thead>
<tr>
<th>PDB Name</th>
<th>Directive for Shares</th>
<th>Directive for Utilization Limit</th>
<th>Memory min</th>
<th>Memory limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALES</td>
<td>3</td>
<td>Unlimited</td>
<td>20</td>
<td>not set</td>
</tr>
<tr>
<td>SERVICES</td>
<td>3</td>
<td>Unlimited</td>
<td>20</td>
<td>not set</td>
</tr>
<tr>
<td>HR</td>
<td>1</td>
<td>70</td>
<td>not set</td>
<td>50</td>
</tr>
</tbody>
</table>

### 6.4 Administering IORM

You can perform various administrative tasks related to I/O Resource Management (IORM).

To perform the tasks, use the `DBMS_RESOURCE_MANAGER` package to define database resource plans on the database servers, and the CellCLI utility to specify the IORM and category plans for each cell.

- **Enabling IORM to Prioritize Latency**
  I/O Resource Management (IORM) is enabled by default and in its default state IORM manages flash cache and flash log.

- **Administering Database Resource Management**
  To set up database resource management, you must use Oracle Database Resource Manager to configure the consumer groups, assign sessions to consumer groups, create a database resource plan, and enable the plan.

- **Administering Interdatabase Resource Management**
  You can configure an interdatabase or category plan for Oracle Exadata Database Machine using the `CellCLI ALTER IORMPLAN` command.

- **Using Flash Cache Attributes in IORM Plans**
  You can use flash cache attributes in the interdatabase plan to guarantee space allocation in Exadata Smart Flash Cache.

- **Managing Flash Cache Quotas for Databases and PDBs**
  IORM enables you to control how you want the flash cache to be shared among different databases and pluggable databases (PDBs).

- **Using PMEM Cache Attributes in IORM Plans**
  You can use PMEM cache attributes in the interdatabase plan to guarantee space allocation in the persistent memory (PMEM) cache.
• **Managing PMEM Cache Quotas for Databases and PDBs**
  I/O Resource Management (IORM) enables you to control how you want the PMEM cache to be shared among different databases and pluggable databases (PDBs).

• **Controlling Access to Flash Cache and Flash Log**
  You can use IORM to manage access to flash cache and flash log for specific databases.

• **Controlling Access to PMEM Cache and PMEM Log**
  You can use IORM to manage access to persistent memory (PMEM) cache and PMEM log for specific databases.

• **Listing an I/O Resource Management Plan**
  You can view the current interdatabase plan for a storage server using the `CellCLI LIST IORMPLAN` command on the storage server.

• **Verifying the Configuration of I/O Resource Management**
  Use this checklist to verify that I/O Resource Management (IORM) is configured correctly.

• **Resetting Default Values in an Interdatabase Plan**
  To reset an interdatabase plan attribute to the default value, set the attribute to an empty string.

**Related Topics**

• **Using the CellCLI Utility**
  You use the Cell Control Command-Line Interface (CellCLI) utility to manage Oracle Exadata System Software.

### 6.4.1 Enabling IORM to Prioritize Latency

I/O Resource Management (IORM) is enabled by default and in its default state IORM manages flash cache and flash log.

IORM also guards against extremely high latencies for log writes, buffer cache reads, and other critical I/Os. Running very heavy I/O loads on hard disks results in high disk I/O latencies. Without IORM, the only way to lower the disk I/O latency is to reduce the overall I/O throughput by reducing smart scans, and other disk intensive I/O operations. To manage these I/O intensive operations, set the `objective` option to specify the optimization mode for IORM. The default `objective` option is `basic`.

When the `objective` option is set to `basic`, the database resource plan maximum utilization limits are not enforced. Database resource plan allocations are only used to guard against extremely high latencies, and there is no consideration for plan conformance.

Smart scan and low priority I/Os to flash devices are queued by IORM in this mode. The critical I/Os usually do not keep the disks busy all the time. During these idle intervals, IORM uses the interdatabase plan allocations to issue the queued smart scan, and low priority I/Os to flash devices. This is done independently of how I/Os to hard disks are processed.

IORM also prioritizes different scan workloads among themselves when no critical OLTP I/Os are being issued. Flash IORM is enabled on hybrid systems with flash cache and Exadata Extreme Flash storage servers. The flash IORM device queue depth varies based on the hardware and flash device type.
For stricter plan conformance, and enforcement of maximum utilization limits, the objective option must be set to something other than basic. The supported IORM objectives are auto, low_latency, balanced, and high_throughput. The recommended objective option is auto which allows IORM to continuously monitor the workloads, and select the best mode based on the active workloads currently on the cells. If the objective option is set to something other than basic, then IORM prioritizes the I/O resources as follows:

1. IORM manages disk I/O if any database has a database resource plan.
2. IORM manages disk I/O if there is a configured interdatabase or category plan.

To deactivate disk I/O prioritization and throttling, set the IORM objective to basic.

Flash IORM behavior cannot be controlled by changing the IORM objective except on Exadata Extreme Flash storage servers. Flash IORM is designed to protect the OLTP latency and does not take the objective into account while protecting the critical I/Os.

On Exadata Extreme Flash storage servers, the IORM objective provides some control over the behavior of flash IORM. The default objective is basic. The objective values of auto and balanced have the same behavior. If the degradation in scan throughput is considered to be too high, the objective can be changed to high_throughput which will increase the scan throughput at the cost of critical I/O latency. The objective may also be changed to low_latency in which case the latency for critical I/Os is very good but there is a significant degradation in scan throughput when both workloads are running concurrently.

6.4.2 Administering Database Resource Management

To set up database resource management, you must use Oracle Database Resource Manager to configure the consumer groups, assign sessions to consumer groups, create a database resource plan, and enable the plan.

- **Setting Up Consumer Groups and Categories**
  Consumer groups and categories are set up with the procedures in the PL/SQL DBMS_RESOURCE_MANAGER package.

- **Assigning Sessions to Consumer Groups**
  You can assign a session to a resource consumer group manually or automatically using consumer group mapping rules.

- **Creating a CDB Plan**
  The CDB plan manages CPU resources on the database servers, and flash cache space and I/O bandwidth on the Exadata storage servers.

- **Creating a Database Plan**
  Database resource plans, also known as intradatabase plans, are created using the PL/SQL procedures DBMS_RESOURCE_MANAGER.CREATE_PLAN() and CREATE_PLAN_DIRECTIVE().

- **Enabling a Database Resource Plan**
  You can manually enable database resource plans by setting the RESOURCE_MANAGER_PLAN parameter. You can automatically enable resource plans by defining an Oracle Scheduler window with a resource plan.

- **Managing Fast File Creation**
  Oracle Exadata System Software features fast file creation, allowing accelerated initialization of data files.
• Managing Data Import
You can control the priority of ETL as well as the amount of I/O resources that ETL consumes using I/O Resource Management (IORM).

• Managing Oracle Recovery Manager Backups and Copies
You can use I/O Resource Management (IORM) to control the resource consumption and priority of RMAN I/Os.

6.4.2.1 Setting Up Consumer Groups and Categories

Consumer groups and categories are set up with the procedures in the PL/SQL DBMS_RESOURCE_MANAGER package.

You can create new consumer groups and categories, or use one of the predefined consumer groups or categories. You do not need to set up categories if you are not planning on using a category plan.

**Note:**
Consumer groups and categories are created in the database and cannot be created explicitly on a cell.

Before running the DBMS_RESOURCE_MANAGER procedures for administering consumer groups and categories, you must first create a pending area. You must have the system privilege ADMINISTER_RESOURCE_MANAGER to run the procedures in the DBMS_RESOURCE_MANAGER PL/SQL package.

The following PL/SQL commands are used with consumer groups and categories:

- **To manage categories:** CREATE_CATEGORY(), DELETE_CATEGORY(), and UPDATE_CATEGORY()

- **To manage consumers groups:** CREATE_CONSUMER_GROUP() and UPDATE_CONSUMER_GROUP()

- **To assign consumer groups to categories:** CREATE_CONSUMER_GROUP() or UPDATE_CONSUMER_GROUP()

In addition to the consumer groups that you set up, the database contains predefined consumer groups. The DBA_RSRC_CONSUMER_GROUPS view displays information about consumer groups, and the DBA_RSRC_CATEGORIES view displays information about categories in the database.

**Example 6-1    Setting Up Consumer Groups and Categories with PL/SQL in the Database**

This example shows how to set up consumer groups and categories in a database.
The MAINTENANCE category is predefined, and is not created in this example.

BEGIN
DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();

DBMS_RESOURCE_MANAGER.CREATE_CATEGORY(
    CATEGORY => 'dss',
    COMMENT => 'DSS consumer groups');
Example 6-2  Consumer Groups and Categories in an Oracle Database

This example shows a query on the `DBA_RSRC_CONSUMER_GROUPS` view.

```sql
SQL> SELECT consumer_group, category FROM DBA_RSRC_CONSUMER_GROUPS
       WHERE consumer_group not like 'ORA%'
       ORDER BY category;
```

<table>
<thead>
<tr>
<th>CONSUMER_GROUP</th>
<th>CATEGORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYS_GROUP</td>
<td>ADMINISTRATIVE</td>
</tr>
<tr>
<td>ETL_GROUP</td>
<td>BATCH</td>
</tr>
<tr>
<td>BATCH_GROUP</td>
<td>BATCH</td>
</tr>
<tr>
<td>DSS_GROUP</td>
<td>BATCH</td>
</tr>
<tr>
<td>CRITICAL_DSS</td>
<td>DSS</td>
</tr>
<tr>
<td>NORMAL_DSS</td>
<td>DSS</td>
</tr>
<tr>
<td>DSS_CRITICAL_GROUP</td>
<td>INTERACTIVE</td>
</tr>
<tr>
<td>INTERACTIVE_GROUP</td>
<td>INTERACTIVE</td>
</tr>
<tr>
<td>ETL</td>
<td>MAINTENANCE</td>
</tr>
<tr>
<td>LOW_GROUP</td>
<td>OTHER</td>
</tr>
<tr>
<td>OTHER_GROUPS</td>
<td>OTHER</td>
</tr>
<tr>
<td>AUTO_TASK_CONSUMER_GROUP</td>
<td>OTHER</td>
</tr>
<tr>
<td>DEFAULT_CONSUMER_GROUP</td>
<td>OTHER</td>
</tr>
</tbody>
</table>

13 rows selected

Related Topics

- Oracle Database Administrator’s Guide
- Oracle Database Reference
6.4.2.2 Assigning Sessions to Consumer Groups

You can assign a session to a resource consumer group manually or automatically using consumer group mapping rules.

For both approaches, you must give explicit permission for a user to switch to a consumer group. In order to control which consumer groups a user can switch to, use the PL/SQL procedure `DBMS_Resource_Manager_Privs.Grant_Switch_Consumer_Group()`. 

The consumer group mapping rules are based on session attributes such as the user name, the name of the service that the session used to connect to the database, and the name of the client program. To create a consumer group mapping rule, use the `SET_CONSUMER_GROUP_MAPPING` procedure, as shown in Example 6-3. Before running the `SET_CONSUMER_GROUP_MAPPING` procedure, you must first create a pending area.

You can also manually switch a session to a particular consumer group, using the PL/SQL `DBMS_Resource_Manager.Switch_Consumer_Group_For_User()` or `SWITCH_CONSUMER_GROUP_FOR_SESS()` procedures.

**Example 6-3  Creating Consumer Group Mapping Rules, Based on Service and User Name**

```sql
BEGIN
    DBMS_SERVICE.CREATE_SERVICE('SALES', 'SALES');
    DBMS_SERVICE.CREATE_SERVICE('AD_HOC', 'AD_HOC');
    DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();
    DBMS_RESOURCE_MANAGER.SET_CONSUMER_GROUP_MAPPING --DBMS_RESOURCE_MANAGER.ORACLE_USER, 'SYS', 'CRITICAL_DSS');
    DBMS_RESOURCE_MANAGER.SET_CONSUMER_GROUP_MAPPING (DBMS_RESOURCE_MANAGER.SERVICE_NAME, 'SALES', 'CRITICAL_DSS');
    DBMS_RESOURCE_MANAGER.SET_CONSUMER_GROUP_MAPPING (DBMS_RESOURCE_MANAGER.SERVICE_NAME, 'AD_HOC', 'NORMAL_DSS');
    DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
    DBMS_RESOURCE_MANAGER.PRIVS.GRANT_SWITCH_CONSUMER_GROUP (GRANTEE_NAME => 'PUBLIC',
        CONSUMER_GROUP => 'CRITICAL_DSS',
        GRANT_OPTION => FALSE);
    DBMS_RESOURCE_MANAGER.PRIVS.GRANT_SWITCH_CONSUMER_GROUP (GRANTEE_NAME => 'PUBLIC',
        CONSUMER_GROUP => 'NORMAL_DSS',
        GRANT_OPTION => FALSE);
END;
/```
6.4.2.3 Creating a CDB Plan

The CDB plan manages CPU resources on the database servers, and flash cache space and I/O bandwidth on the Exadata storage servers.

CDB plans are created using the PL/SQL procedures

`DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN()` and `CREATE_CDB_PLAN_DIRECTIVE()`.

The CDB plan can only be configured from the root PDB.

Example 6-4  Using a CDB Plan to Distribute Resources Between PDBs

This example shows how to distribute resources between three PDBs named SALES, SERVICES and HR. SALES and SERVICES have higher priority and get three shares each compared to one share for HR. The limit on the HR PDB is set to 70% maximum utilization limit.

```sql
BEGIN
    DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();

    DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN(
        plan     => 'NEWCDB_PLAN ',
        comment  => 'CDB resource plan for newcdb');

    DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
        plan      => 'NEWCDB_PLAN ',
        pluggable_database => 'SALESPDB',
        shares       => 3,
        memory_min   => 20,
        utilization_limit => 100);

    DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
        plan      => 'NEWCDB_PLAN ',
        pluggable_database => 'SERVICESPDB',
        shares       => 3,
        memory_min   => 20,
        memory_limit => 75);

    DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
        plan      => 'NEWCDB_PLAN ',
        pluggable_database => 'HRPDB',
        shares       => 1,
        memory_limit => 50,
        utilization_limit => 70);

    DBMS_RESOURCE_MANAGER.VALIDATE_PENDING_AREA();
    DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
```
6.4.2.4 Creating a Database Plan

Database resource plans, also known as **intradatabase plans**, are created using the PL/SQL procedures `DBMS_RESOURCE_MANAGER.CREATE_PLAN()` and `CREATE_PLAN_DIRECTIVE()`.

You must always begin resource plan creations or updates with the PL/SQL procedure `CREATE_PENDING_AREA()` and complete them with the PL/SQL procedure `SUBMIT_PENDING_AREA()`. You must also include a directive for `OTHER_GROUPS`, which includes all sessions that are not explicitly mapped to a consumer group.

You must have the system privilege `ADMINISTER_RESOURCE_MANAGER` to run the procedures in the `DBMS_RESOURCE_MANAGER` PL/SQL package. This resource plan manages both CPU resources on database instances and I/O resources on the cells.

**Example 6-5  Sharing Resources Across Applications**

In this example, assume you have multiple applications sharing a database where the I/O resources should be divided across the applications using a particular ratio. For example, there are three applications named SALES, FINANCE, and MARKETING. You would like the I/O resources to be allocated as 60 percent, 25 percent, and 10 percent, respectively, with the remaining 5 percent allocated to any sessions that do not map into these consumer groups. In this scenario, you would create a consumer group for each application, and then create a single-level resource plan and specify the percentage of I/O resources for each consumer group. This allocation is actually the minimum I/O resources that the consumer group can use. If a consumer group does not use its allocation, then it is redistributed to the other consumer groups in the ratio specified by the plan. You can specify the allocations using the `MGMT_P1` parameter.

```sql
BEGIN
    DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();
    DBMS_RESOURCE_MANAGER.CREATE_PLAN('DAYTIME_PLAN', 'Resource plan for managing all applications between 9 am and 5 pm');
    DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('SALES', 'Sales App');
    DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('FINANCE', 'Finance App');
    DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('MARKETING', 'Marketing App');
    DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'SALES', 'Allocation for SALES', MGMT_P1 => 60);
    DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'FINANCE', 'Allocation for FINANCE', MGMT_P1 => 25);
    DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'MARKETING', 'Allocation for MARKETING', MGMT_P1 => 10);
    DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'OTHER_GROUPS', 'Allocation for default group', MGMT_P1 => 5);
    DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
END;
```
Example 6-6    Sharing Resources Across Workloads

In this example, assume you want to prioritize one workload over another. For example, suppose that you load data into your data warehouse while also servicing queries, and you want to always prioritize the queries over the data load. For this scenario, you would create two consumer groups for queries (reporting and ad-hoc) and one consumer group for data load. You would like to share the I/O resources between the two query consumer groups using a 75/25 ratio. In addition, you would like to issue I/Os for data load only if the query consumer groups do not use all of their allocation. You can use resource plan levels to specify the allocation priorities.

BEGIN
DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();
DBMS_RESOURCE_MANAGER.CREATE_PLAN('DAYTIME_PLAN', 'Resource plan for prioritizing queries between 9 am and 5 pm');
DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('REPORT_QUERIES', 'Report Queries');
DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('AD-HOC_QUERIES', 'Ad-Hoc Queries');
DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('DATA_LOAD', 'Data Load');

DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'REPORT_QUERIES', 'Allocation for REPORT_QUERIES', MGMT_P1 => 75);
DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'AD-HOC_QUERIES', 'Allocation for AD-HOC_QUERIES', MGMT_P1 => 25);
DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'DATA_LOAD', 'Allocation for DATA_LOAD', MGMT_P2 => 100);
DBMS_RESOURCE_MANAGER.CREATE_PLAN_DIRECTIVE('DAYTIME_PLAN', 'OTHER_GROUPS', 'Allocation for default group', MGMT_P3 => 100);
DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
END;
/

Related Topics

• Oracle Database Administrator's Guide

6.4.2.5 Enabling a Database Resource Plan

You can manually enable database resource plans by setting the RESOURCE_MANAGER_PLAN parameter. You can automatically enable resource plans by defining an Oracle Scheduler window with a resource plan.

When the Oracle Scheduler window opens, the resource plan is enabled. When the Oracle Scheduler window closes, the resource plan is disabled.

When a resource plan is enabled, the database alerts all cells about this event and provides the resource plan. The database also alerts all cells when a resource plan
is disabled. Because only one resource plan can be active for any database, you are required to enable the same resource plan on all instances of a database. If no database resource plan is enabled for a database, then all I/O requests are treated equally.

Related Topics
- Oracle Database Administrator’s Guide

6.4.2.6 Managing Fast File Creation

Oracle Exadata System Software features fast file creation, allowing accelerated initialization of data files.

This feature automatically runs whenever you create a new tablespace, add a data file to an existing tablespace, or autoextend an existing tablespace. Oracle Exadata System Software can initialize files very quickly because it issues many concurrent I/O requests. However, these concurrent I/O requests create a heavy load that can interfere with performance-critical queries.

Using I/O Resource Management (IORM), you can control the priority of fast file creations for creating a new tablespace or adding a data file to an existing tablespace. These operations are run under the `FASTFILECRE` function. By default, the `FASTFILECRE` function is mapped to a hidden consumer group that has lower priority than all consumer group and background I/Os. If you choose to increase the priority, and thereby performance, of file creations, add a mapping rule based on the mapping attribute `DBMS_RESOUCRE_MANAGER.ORACLE_FUNCTION`, and mapping value `FASTFILECRE`.

Because autoextending an existing tablespace is a brief and often time-critical operation, you cannot modify its priority using IORM.

Example 6-7  Managing Fast File Creation

```
BEGIN
DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();
DBMS_RESOURCE_MANAGER.CREATE_CONSUMER_GROUP('MAINTENANCE_GROUP', 'Maintenance activity');
DBMS_RESOURCE_MANAGER.SET_CONSUMER_GROUP_MAPPING(DBMS_RESOURCE_MANAGER.ORACLE_FUNCTION, 'FASTFILECRE', 'MAINTENANCE_GROUP');
DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
END;
/
```

6.4.2.7 Managing Data Import

You can control the priority of ETL as well as the amount of I/O resources that ETL consumes using I/O Resource Management (IORM).

Data import, or extract, transform, load (ETL), is an important part of maintaining a data warehouse. In some cases, ETL is extremely critical to performance because reports or queries cannot be run until the data has been loaded. In these cases, ETL should be prioritized above all other queries. In other cases, ETL is a low-priority background activity that only needs to be prioritized in the rare event that it does not complete by a certain time.
To manage ETL, do the following:

- Map the ETL sessions to the `ETL_GROUP` consumer group.
  
  The mapping rules for ETL are typically based on user name or client program name. Data pump is run under the `DATALOAD` function. By default, the `DATALOAD` function is mapped to the `ETL_GROUP` consumer group.

- Include the `ETL_GROUP` group in your resource plans.

To import non-compressed data as compressed data,

**Example 6-8  Mapping a Program to the ETL_GROUP Consumer Group**

This example shows how to map a program to the `ETL_GROUP` consumer group.

```sql
BEGIN
  DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();
  DBMS_RESOURCE_MANAGER.SET_CONSUMER_GROUP_MAPPING
    (DBMS_RESOURCE_MANAGER.CLIENT_PROGRAM, 'SQLLDR', 'ETL_GROUP');
  DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
END;
/
```

**Importing Non-compressed Data as Compressed Data**

Non-compressed data can be imported as compressed data when using the `TRANSFORM:SEGMENT_ATTRIBUTES=n` option, and the target tablespace has been configured to create new tables as Exadata Hybrid Columnar Compression tables by default.

**Related Topics**

- TRANSFORM

### 6.4.2.8 Managing Oracle Recovery Manager Backups and Copies

You can use I/O Resource Management (IORM) to control the resource consumption and priority of RMAN I/Os.

Backups are an I/O intensive operation. You can control the rate of Oracle Recovery Manager (RMAN) I/Os by configuring the number of channels. You can use IORM for a greater degree of control over the resource consumption and priority of RMAN I/Os. For example, you can map RMAN to a low priority consumer group. If the Oracle Exadata Storage Server is busy, then the RMAN operations run very slowly and not interfere with the other database operations. However, whenever the Oracle Exadata Storage Server is not fully utilized, then IORM schedules the RMAN I/Os, allowing it to consume the unutilized bandwidth.

RMAN backups run under the `BACKUP` function. RMAN copies run under the `COPY` function. By default, both the `BACKUP` and `COPY` functions are mapped to the `BATCH_GROUP` consumer group. You can remap these functions to any other consumer group, as shown in the following example.
Example 6-9 Using Consumer Groups to Manage Resources

This example shows how to map the \texttt{BACKUP} function to the \texttt{BATCH\_GROUP} consumer group and the \texttt{COPY} function to the \texttt{MAINTENANCE\_GROUP} consumer group.

```sql
BEGIN
  DBMS\_RESOURCE\_MANAGER\_.CREATE\_PENDING\_AREA();
  DBMS\_RESOURCE\_MANAGER\_.SET\_CONSUMER\_GROUP\_MAPPING(DBMS\_RESOURCE\_MANAGER\_.ORACLE\_FUNCTION, 'BACKUP', 'BATCH\_GROUP');
  DBMS\_RESOURCE\_MANAGER\_.SET\_CONSUMER\_GROUP\_MAPPING(DBMS\_RESOURCE\_MANAGER\_.ORACLE\_FUNCTION, 'COPY', 'MAINTENANCE\_GROUP');
  DBMS\_RESOURCE\_MANAGER\_.SUBMIT\_PENDING\_AREA();
END;
/
```

6.4.3 Administering Interdatabase Resource Management

You can configure an interdatabase or category plan for Oracle Exadata Database Machine using the CellCLI \texttt{ALTER\ IORMPLAN} command.

- **How to Create an Interdatabase or Category Plan**
  You can run the \texttt{ALTER\ IORMPLAN} commands in CellCLI.

- **Required Settings for an Interdatabase or Category Plan**
  Each category or interdatabase plan contain the required settings.

- **Optional Settings for an Interdatabase or Category Plan**
  After specifying the required directives, you can include additional directives to customize your plan.

6.4.3.1 How to Create an Interdatabase or Category Plan

You can run the \texttt{ALTER\ IORMPLAN} commands in CellCLI.

You can also put the \texttt{ALTER\ IORMPLAN} commands in a text file, such as \texttt{alter\_iorm}, then run the commands using the text file with the CellCLI \texttt{START\ alter\_iorm} command. The plan name is automatically set to \texttt{cellname\_IORMPLAN}.

The \texttt{catPlan} parameter specifies the category plan. The \texttt{dbPlan} parameter specifies the interdatabase plan. To manage the I/O resources using a category plan, interdatabase plan or database resource plan, set the objective.

The default for the \texttt{objective} option is \texttt{basic}, and I/O Resource Management (IORM) manages I/O resources based on the configured resource plans. The \texttt{objective} option can be set to \texttt{basic}, \texttt{auto}, \texttt{low\_latency}, \texttt{balanced}, or \texttt{high\_throughput}.

If a container database (CDB) plan specifies the \texttt{memory\_min} and \texttt{memory\_limit} attributes for pluggable database (PDB), then it is computed as a percentage of the \texttt{flashcachemin} and \texttt{flashcachelimit} or \texttt{flashcachesize} values for the database or the total flash cache size when nothing is specified. The \texttt{memory\_min} and \texttt{memory\_limit} are also applied to PMEM cache and it is also computed as a percentage of the \texttt{pmemcachemin} and \texttt{pmemcachelimit} or \texttt{pmemcachesize} values for the database or the total PMEM cache size when nothing is specified.
6.4.3.2 Required Settings for an Interdatabase or Category Plan

Each category or interdatabase plan contain the required settings.

For interdatabase plan directives to be valid, you must provide either a level and an allocation, share, maximum utilization limit, or a flash cache quota directive. A directive without an allocation, share or maximum utilization limit is invalid. A directive that specifies both allocation and share is also invalid.

When specifying catPlan or dbPlan, a directive with name=other is required. For category plans, the other directive provides an allocation for all active consumer groups whose category is not specified in the category plan. For interdatabase plans, the other directive provides an allocation for all databases that are using Oracle Exadata System Software, but are not explicitly specified in the interdatabase plan. If the other directive is not specified, then the CellCLI utility returns an error.

6.4.3.2.1 Using Allocations to Specify Resources

You can prioritize the I/O requests from the four example databases by specifying an interdatabase plan.

Consider the example of four databases sharing the same storage. The four databases are:

- An OLTP production database, named PROD
- A test database, named PROD_TEST
- A development database, named PROD_DEV
- A data warehouse database, named DW

An OLTP production database typically issues small I/O requests, and low latency for these requests is the critical requirement. A data warehouse issues large numbers of large I/O requests and is more sensitive to the I/O throughput than the latency of each individual I/O request. Without any I/O resource management, the large number of I/O requests issued by the DW database could overwhelm the storage subsystem and increase the latency of the I/O requests issued by the PROD database. Additionally, the I/O requests issued by the test and development databases, PROD_TEST and PROD_DEV, could adversely affect the performance of the PROD and the DW databases.

The interdatabase plan has the following characteristics:
• The OLTP database PROD gets 80 percent of the I/O resources with the highest priority level.

• The DW database gets 20 percent of the remaining I/O resources and 80 percent of PROD's unused allocation.

• The PROD_TEST and PROD_DEV databases and the OTHER consumer group get any unused I/O in the amount of 50 percent, 40 percent, and 10 percent respectively.

An example of this interdatabase plan is shown in the following table.

### Table 6-4 Interdatabase Plan Using Allocations

<table>
<thead>
<tr>
<th>Database Name</th>
<th>Level 1 (%)</th>
<th>Level 2 (%)</th>
<th>Level 3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROD</td>
<td>80</td>
<td>not set</td>
<td>not set</td>
</tr>
<tr>
<td>DW</td>
<td>not set</td>
<td>80</td>
<td>not set</td>
</tr>
<tr>
<td>PROD_TEST</td>
<td>not set</td>
<td>not set</td>
<td>50</td>
</tr>
<tr>
<td>PROD_DEV</td>
<td>not set</td>
<td>not set</td>
<td>40</td>
</tr>
<tr>
<td>OTHER</td>
<td>not set</td>
<td>not set</td>
<td>10</td>
</tr>
</tbody>
</table>

If the PROD_TEST and PROD_DEV databases put out an inordinately large I/O load, then the performance of databases PROD and DW are not affected. Also, if database DW issues large amounts of I/O, then the performance of the PROD database is not adversely affected.

### Example 6-10 ALTER IORMPLAN Syntax for Allocations

This example shows the command to create the interdatabase plan described previously in this topic.

```
CellCLI> ALTER IORMPLAN
dbPlan=(
    (name=prod, level=1, allocation=80),
    (name=dw, level=2, allocation=80),
    (name=prod_test, level=3, allocation=50),
    (name=prod_dev, level=3, allocation=40),
    (name=other, level=3, allocation=10))
```

### Example 6-11 Configuring an Interdatabase Plan Using Allocations

This example shows how to configure an interdatabase plan using allocation percentages for a cell.

```
CellCLI> ALTER IORMPLAN
catPlan=((name=administrative, level=1, allocation=80),
    (name=interactive, level=2, allocation=90),
    (name=batch, level=3, allocation=80),
    (name=maintenance, level=4, allocation=50),
    (name=other, level=4, allocation=50))

dbplan=((name=sales_prod, share=8, role=primary),
    (name=sales_prod, share=1, limit=50, role=standby),
    (name=sales_test, share=1, limit=25),
    (name=default, share=2))
```
6.4.3.2.2 Using Shares to Specify Resources

You can prioritize the I/O requests from the four example databases by assigning a share to each database in an interdatabase plan.

Consider the example of four databases sharing the same storage. The four databases are:

- An OLTP production database, named `PROD`
- A test database, named `PROD_TEST`
- A development database, named `PROD_DEV`
- A data warehouse database, named `DW`

An OLTP production database typically issues small I/O requests, and low latency for these requests is the critical requirement. A data warehouse issues large numbers of large I/O requests and is more sensitive to the I/O throughput than the latency of each individual I/O request. Without any I/O resource management, the large number of I/O requests issued by the `DW` database could overwhelm the storage subsystem and increase the latency of the I/O requests issued by the `PROD` database. Additionally, the I/O requests issued by the test and development databases, `PROD_TEST` and `PROD_DEV`, could adversely affect the performance of the `PROD` and `DW` databases.

Because the OLTP database `PROD` is most critical, you can specify 16 shares for the `PROD` database. The `DW` database gets 4 shares, `PROD_TEST` gets 2 shares, and `PROD_DEV` gets 1 share. This plan ensures that the `PROD` database is 4 times more likely than the `DW` database to issue an I/O.

If `PROD` does not have any I/Os to issue, then the other databases are picked based on their shares.

**Example 6-12  ALTER IORMPLAN Syntax for Shares**

This example shows the command to create the interdatabase plan by assigning shares to each database.

```
CellCLI> ALTER IORMPLAN
   dbPlan=(
          (name=prod, share=16),
          (name=dw, share=4),
          (name=prod_test, share=2),
          (name=prod_dev, share=1))
```

**Example 6-13  Configuring an Interdatabase Plan Using Shares**

This example shows how to configure an interdatabase plan using shares for a cell.

```
ALTER IORMPLAN
   dbPlan=(
          (name=dev01, share=1, limit=50, flashLog=off),
          (name=dev02, share=1, limit=25, flashCache=off),
          (name=other, share=4))
```
6.4.3.3 Optional Settings for an Interdatabase or Category Plan

After specifying the required directives, you can include additional directives to customize your plan.

When configuring an interdatabase plan, catPlan and dbPlan are optional. If catPlan is not specified, then IORM between categories is not enabled. Similarly, if dbPlan is not specified, then IORM between databases is not enabled.

The optional directives include:

- **role Attribute for IORM Plans**
  The role attribute allows a different allocation to be specified, based on whether the database has the Oracle Data Guard primary or standby role.

- **asmcluster Attribute for IORM Plans**
  You can use the asmcluster attribute to uniquely identify databases with the same DB_UNIQUE_NAME.

### 6.4.3.3.1 role Attribute for IORM Plans

The role attribute allows a different allocation to be specified, based on whether the database has the Oracle Data Guard primary or standby role.

By default, all interdatabase plan allocations apply when the database is in either role. If you want the allocation to apply only when the database is in the primary role, then include role=primary in the plan. Similarly, if you want the allocation to apply only when the database is in the standby role, then include role=standby in the plan.

```sql
ALTER IORMPLAN
- dbPlan=((name=prod, share=8, role=primary),
  (name=prod, share=1, limit=25, role=standby),
  (name=default, share=2))
```

**Related Topics**

- Configuring Resource Manager for Oracle Active Data Guard (My Oracle Support Doc ID 1930540.1)

### 6.4.3.3.2 asmcluster Attribute for IORM Plans

You can use the asmcluster attribute to uniquely identify databases with the same DB_UNIQUE_NAME.

If you have databases with the same DB_UNIQUE_NAME but associated with different Oracle ASM clusters, then, starting with Oracle Exadata System Software release 19.1.0, you can use the asmcluster attribute to uniquely identify the databases.

The databases must be clients of different Oracle ASM clusters, and you must have already configured ASM-scoped security.

```sql
ALTER IORMPLAN
- dbplan=((name=pdb1, share=4, flashcachemin=5G, asmcluster=asm1),
  (name=pdb1, share=2, limit=80, asmcluster=asm2),
```
6.4.4 Using Flash Cache Attributes in IORM Plans

You can use flash cache attributes in the interdatabase plan to guarantee space allocation in Exadata Smart Flash Cache.

Using flash cache attributes, IORM can reserve space for critical databases while preventing less important or rogue databases from consuming the entire flash cache. These attributes can only be specified in an interdatabase plan, and are configured using the CellCLI utility.

You can use the following flash cache attributes in the dbPlan to set hard or soft limits for flash cache resources. A hard limit means the database cannot exceed its quota even when the memory cache is not full. A soft limit means that the specified limit can be exceeded if there are available resources.

- **flashCacheMin** — Specifies the minimum amount of flash cache space that is guaranteed for a database, even if the blocks are cold. This is a hard limit. Because flashCacheMin is a guaranteed reservation, the sum of flashCacheMin across all directives should be less than the size of the flash cache to ensure that each database gets its respective quota.

- **flashCacheLimit** — Specifies the soft maximum amount of flash cache space that is available to a database. If the flash cache is not full, a database can exceed the flashCacheLimit value.

- **flashCacheSize** — Specifies the hard maximum amount of flash cache space that is available to a database. A database is not allowed to exceed the flashCacheSize value at any time.

However, if you set flashCacheSize to a value that is lower than the current space occupied by the database, then starting with Oracle Exadata System Software release 20.1.0 excess data is proactively cleared from the cache. Previously, excess data was only removed when overwritten by other data. Starting with Oracle Exadata System Software release 19.2.0, flashCacheSize is not a guaranteed reservation if the sum of the flashCacheSize across all directives is more than the size of the flash cache. In this case, you can also specify flashCacheMin to define a guaranteed minimum quota.

**Example 6-14 Configuring an Interdatabase Plan with Flash Cache Attributes**

This example shows how to create an interdatabase plan with flash cache attributes. In the example, the sales and test databases are guaranteed an amount of space in the flash cache by using the flashCacheSize parameter. But, the databases cannot exceed the specified allocation, even if the flash cache has free space.
The `finance` and `dev` databases use `flashCacheMin` for guaranteed minimum quotas. These databases can also exceed the specified `flashCacheLimit` size when there is free space in the flash cache.

```
ALTER IORMPLAN
  dbplan=((name=sales, share=8,
    flashCacheSize=10G),
    (name=finance, share=8, flashCacheLimit=10G,
      flashCacheMin=2G),
    (name=dev, share=2, flashCacheLimit=4G,
      flashCacheMin=1G),
    (name=test, share=1, limit=10, flashCacheSize=1G))
```

**Example 6-15  Configuring an Interdatabase Plan with Flash Cache Attributes**

Starting with Oracle Exadata System Software release 19.3.0, you can specify both `flashCacheMin` and `flashCacheSize` for the same target.

```
ALTER IORMPLAN
  dbplan=((name=sales, share=8,
    flashCacheMin=3G,
    flashCacheSize=10G),
    (name=finance, share=8, flashCacheLimit=10G,
      flashCacheMin=2G),
    (name=dev, share=2, flashCacheLimit=4G,
      flashCacheMin=1G),
    (name=test, share=1, limit=10, flashCacheSize=1G))
```

### 6.4.5 Managing Flash Cache Quotas for Databases and PDBs

IORM enables you to control how you want the flash cache to be shared among different databases and pluggable databases (PDBs).

This can be done using just the CDB resource plan or both the CDB plan and the I/O Resource Management (IORM) interdatabase plan.

Consider a CDB resource plan that specifies `memory_min` and `memory_limit` for the 3 PDBs mentioned in the plan. Note the following:

- These values are specified in percentages and range from 0 and 100. Because over-provisioning is supported, the sum of the percentages is not restricted to 100%. If the sum of these values is greater than 100%, then the values are normalized down to 100%.
- If `memory_min` is not specified, then it defaults to 0.
- If `memory_limit` is not specified, then it defaults to 100.
- For `CDB$ROOT`, there is a 5% `memory_limit` value.

The following example shows how to create an interdatabase plan for 3 PDBs. The sum of the `memory_min` values is 40%, and the sum of the `memory_limit` values is 175%, which must be normalized. If an interdatabase plan is not specified, then these percentages apply to the entire size of the flash cache. If an interdatabase plan is specified, then the quotas for the PDBs are computed as a percentage...
of the `memory_min` and `memory_limit` values for each database as specified in the interdatabase plan directive.

BEGIN
DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN(
    plan       => 'NEWCDB_PLAN',
    comment    => 'CDB resource plan for newcdb');

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
    plan => 'NEWCDB_PLAN',
    pluggable_database => 'SALESPDB',
    memory_min => 20);

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
    plan => 'NEWCDB_PLAN',
    pluggable_database => 'SERVICESPDB',
    memory_min => 20,
    memory_limit => 50);

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
    plan => 'NEWCDB_PLAN',
    pluggable_database => 'HRPDB',
    memory_limit => 25);

DBMS_RESOURCE_MANAGER.VALIDATE_PENDING_AREA();
DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
END;
/

In the example above, if no interdatabase plan is specified and the size of the flash cache is 10 GB, then the following table shows the breakdown of the quotas after normalization of the limit (because the sum of the `memory_limit` values is greater than 100%). If we end up with a minimum value that is greater than the limit, we reduce the minimum value to make it equal to the limit.

<table>
<thead>
<tr>
<th>PDB</th>
<th>Flash Cache Min</th>
<th>FC Soft Limit</th>
<th>Normalized Soft Limit</th>
<th>FC Hard Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALESPDB</td>
<td>20% = 2 GB</td>
<td>100 (default)</td>
<td>100 / 175 = 5.7 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>SERVICESPDB</td>
<td>20% = 2 GB</td>
<td>50</td>
<td>50 / 175 = 2.85 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>HRPDB</td>
<td>0</td>
<td>25</td>
<td>25 / 175 = 1.4 GB</td>
<td>n/a</td>
</tr>
</tbody>
</table>

The following example shows an interdatabase plan running on a system where the flash cache size is 50 GB and PMEM cache size is 10 GB.

```sql
ALTER IORMPLAN dbplan=
  ((name=newcdb, share=8, pmemCacheSize= 2G, flashCacheSize=10G),
   (name=finance, share=8, pmemCacheMin= 1G, pmemCacheLimit= 2G,
    flashCacheLimit=10G, flashCacheMin=2G),
   (name=dev, share=2, pmemCacheMin= 100M, pmemCacheLimit= 1G,
```
In addition to the NEWCDB CDB, three other databases (finance, dev, and test) share the same storage servers. Flash cache quotas are only enforced if the directives specify the flashcachesize, flashcachelimit, or flashcachemin attribute. The flashcachesize attribute specifies a hard limit and cannot be specified with flashcachelimit in the same directive. The database test does not specify any flash cache directive; so that database and its PDBs (if any exist) are not managed for any flash cache quotas.

If the CDB has flashcachesize specified, then the memory_min value from the CDB resource plan is ignored, and the memory_limit value is normalized and used to compute the flash cache sizes for the different PDBs. Because the newcdb CDB has flashcachesize specified in the example above, the memory_min values are ignored. The flashcachesize is computed after normalizing the memory_limit values, as was shown previously. The only difference is that this will be a hard limit since the CDB has specified the flashcachesize directive.

<table>
<thead>
<tr>
<th>PDB</th>
<th>Flash Cache Min</th>
<th>FC Hard Limit</th>
<th>Normalized Hard Limit</th>
<th>FC Soft Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALESPDB</td>
<td>0</td>
<td>100 (default)</td>
<td>100 / 175 = 5.71 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>SERVICESPDB</td>
<td>0</td>
<td>50</td>
<td>50 / 175 = 2.86 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>HRPDB</td>
<td>0</td>
<td>25</td>
<td>25 / 175 = 1.43 GB</td>
<td>n/a</td>
</tr>
</tbody>
</table>

For non-CDB databases, the flashcachesize, flashcachemin, and flashcachelimit values are specified in absolute terms and no additional normalization is required. Because flashcachemin is a guaranteed reservation, the sum of flashcachemin across all the directives should be less than the total size of the flash cache.

6.4.6 Using PMEM Cache Attributes in IORM Plans

You can use PMEM cache attributes in the interdatabase plan to guarantee space allocation in the persistent memory (PMEM) cache.

Using PMEM cache attributes, IORM can reserve space for critical databases while preventing less important or rogue databases from consuming the entire PMEM cache. These attributes can only be specified in an interdatabase plan, and are configured using the CellCLI utility.

You can use the following PMEM cache attributes in the dbPlan to set hard or soft limits for PMEM cache resources. A hard limit means the database cannot exceed its quota even when the memory cache is not full. A soft limit means that the specified limit can be exceeded if there are available resources.

- **pmemCacheMin** — Specifies the minimum amount of PMEM cache space that is guaranteed for a database, even if the blocks are cold. This is a hard limit.
Because `pmemCacheMin` is a guaranteed reservation, the sum of `pmemCacheMin` across all directives should be less than the size of the PMEM cache to ensure that each database gets its respective quota.

- `pmemCacheLimit` — Specifies the soft maximum amount of PMEM cache space that is available to a database. If the PMEM cache is not full, a database can exceed the `pmemCacheLimit` value.

- `pmemCacheSize` — Specifies the hard maximum amount of PMEM cache space that is available to a database. A database is not allowed to exceed the `pmemCacheSize` value at any time.

However, if you set `pmemCacheSize` to a value that is lower than the current space occupied by the database, then starting with Oracle Exadata System Software release 20.1.0 excess data is proactively cleared from the cache. Previously, excess data was only removed when overwritten by other data.

`pmemCacheSize` is not a guaranteed reservation if the sum of the `pmemCacheSize` across all directives is more than the size of the PMEM cache. In this case, you can also specify `pmemCacheMin` to define a guaranteed minimum quota.

**Example 6-16  Configuring an Interdatabase Plan with PMEM Cache Attributes**

This example shows how to create an interdatabase plan with pmem cache attributes. In the example, the `sales` and `test` databases are guaranteed an amount of space in the PMEM cache by using the `pmemCacheSize` parameter. But, the databases cannot exceed the specified allocation, even if the PMEM cache has free space.

The `finc` and `dev` databases use `pmemCacheMin` for guaranteed minimum quotas. These databases can also exceed the specified `pmemCacheLimit` size when there is free space in the PMEM cache.

The example plan also contains various flash cache attributes.

```
ALTER IORMPLAN dbplan=                                            -
    ((name=sales, share=8, pmemCacheSize= 2G, flashCacheSize=10G), -
    (name=finc, share=8, pmemCacheMin= 1G, pmemCacheLimit= 2G, -
     flashCacheLimit=10G, flashCacheMin=2G), -
    (name=dev, share=2, pmemCacheMin= 500M, pmemCacheLimit= 1G, -
     flashCacheLimit=4G, flashCacheMin=1G), -
    (name=test, share=1, limit=10, pmemCacheSize= 200M))
```

### 6.4.7 Managing PMEM Cache Quotas for Databases and PDBs

I/O Resource Management (IORM) enables you to control how you want the PMEM cache to be shared among different databases and pluggable databases (PDBs).

Consider a CDB resource plan that specifies `memory_min` and `memory_limit` for the 3 PDBs mentioned in the plan. Note the following:

- These values are specified in percentages and range from 0 and 100. Because over-provisioning is supported, the sum of the percentages is not restricted to 100%. If the sum of these values is greater than 100%, then the values are normalized down to 100%.

- If `memory_min` is not specified, then it defaults to 0.

- If `memory_limit` is not specified, then it defaults to 100.
For CDB$ROOT, there is a 5% memory_limit value.

The following example shows how to create an interdatabase plan for 3 PDBs. The sum of the memory_min values is 40%, and the sum of the memory_limit values is 175%, which must be normalized. If an interdatabase plan is not specified, then these percentages apply to the entire size of the PMEM cache. If an interdatabase plan is specified, then the quotas for the PDBs are computed as a percentage of the memory_min and memory_limit values for the database as specified in the interdatabase plan directive.

BEGIN
DBMS_RESOURCE_MANAGER.CREATE_PENDING_AREA();

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN(
    plan => 'NEWCDB_PLAN ',
    comment => 'CDB resource plan for newcdb');

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
    plan => 'NEWCDB_PLAN ',
    pluggable_database => 'SALESPDB',
    memory_min => 20);

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
    plan => 'NEWCDB_PLAN ',
    pluggable_database => 'SERVICESPDB',
    memory_min => 20,
    memory_limit => 50);

DBMS_RESOURCE_MANAGER.CREATE_CDB_PLAN_DIRECTIVE(
    plan => 'NEWCDB_PLAN ',
    pluggable_database => 'HRPDB',
    memory_limit => 25);

DBMS_RESOURCE_MANAGER.VALIDATE_PENDING_AREA();
DBMS_RESOURCE_MANAGER.SUBMIT_PENDING_AREA();
END;
/

In the example above, if no interdatabase plan is specified and the size of the PMEM cache is 10 GB, then the following table shows the breakdown of the quotas after normalization of the limit where the sum of the memory_limit values is greater than 100%. If we end up with a minimum value that is greater than the limit, we reduce the minimum value to make it equal to the limit.

Table 6-7  Case 3: PDB PMEM Cache Limits with No Interdatabase Plan

<table>
<thead>
<tr>
<th>PDB</th>
<th>PMEM Cache Min</th>
<th>PMEM Soft Limit</th>
<th>Normalized Soft Limit</th>
<th>PMEM Hard Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALESPDB</td>
<td>20% = 2 GB</td>
<td>100 (default)</td>
<td>100 / 175 = 5.7 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>SERVICESPDB</td>
<td>20% = 2 GB</td>
<td>50</td>
<td>50 / 175 = 2.85 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>HRPDB</td>
<td>0</td>
<td>25</td>
<td>25 / 175 = 1.4 GB</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Next, consider an interdatabase plan, as shown below, running on a system where the flash cache size is 50 GB and PMEM cache size is 10 GB.

```
ALTER IORMPLAN dbplan=
    ((name=newcdb, share=8, pmemCacheSize= 2G, flashCacheSize=10G),
     (name=finance, share=8, pmemCacheMin= 1G, pmemCacheLimit= 2G,
      flashCacheLimit=10G, flashCacheMin=2G),
     (name=dev, share=2, pmemCacheMin= 100M, pmemCacheLimit= 1G,
      flashCacheLimit=4G, flashCacheMin=1G),
     (name=test, share=1, limit=10))
```

In addition to the newcdb CDB, three other databases (finance, dev, and test) share the same storage servers. PMEM cache quotas are only enforced if the directives specify the `pmemcachemin`, `pmemcachelimit` and `pmemcachesize` attributes. `pmemcachesize` is a hard limit and it cannot be specified with `pmemcachelimit` in the same directive. Database test does not specify any PMEM cache directive, so that database and its PDBs (if any exist) are not managed for any PMEM cache quotas.

If the CDB has `pmemcachesize` specified, then the `memory_min` value from the CDB resource plan is ignored, and the `memory_limit` value is normalized and used to compute the PMEM cache sizes for the different PDBs. Because the newcdb CDB has `pmemcachesize` specified, the `memory_min` values are ignored. The `pmemcachesize` is computed after normalizing the `memory_limit` values, as was shown previously. The only difference is that this will be a hard limit since the CDB has specified the `pmemcachesize` directive.

<table>
<thead>
<tr>
<th>PDB</th>
<th>PMEM Cache Min</th>
<th>PMEM Hard Limit</th>
<th>Normalized Hard Limit</th>
<th>PMEM Soft Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALESPDB</td>
<td>0</td>
<td>100 (default)</td>
<td>100 / 175 = 1.14 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>SERVICESPDB</td>
<td>0</td>
<td>50</td>
<td>50 / 175 = 0.57 GB</td>
<td>n/a</td>
</tr>
<tr>
<td>HRPDB</td>
<td>0</td>
<td>25</td>
<td>25 / 175 = 0.28 GB</td>
<td>n/a</td>
</tr>
</tbody>
</table>

For non-CDB databases, the `pmemcachesize`, `pmemcachemin`, and `pmemcachelimit` values are specified in absolute terms and no additional normalization is required. Because `pmemcachemin` is a guaranteed reservation, the sum of `pmemcachemin` across all the directives should be less than the total size of the PMEM cache.

### 6.4.8 Controlling Access to Flash Cache and Flash Log

You can use IORM to manage access to flash cache and flash log for specific databases.

You can use the `flashcache` and `flashlog` directives in the interdatabase IORM plan to control access to flash cache and flash log.
The following example shows how to disable flash cache and flash log for the `prod_test` and `prod_dev` databases. In the example, flash log is also disabled for the `dw_test` database.

```
CellCLI> ALTER IORMPLAN
   -
   dbPlan=(
      -
         (name=prod, flashcache=on, flashlog=on),
         (name=dw, flashcache=on, flashlog=on),
         (name=prod_test, flashcache=off, flashlog=off),
         (name=prod_dev, flashcache=off, flashlog=off),
         (name=dw_test, flashcache=on, flashlog=off))
```

You can also use these directives in conjunction with other directives. For example:

```
CellCLI> ALTER IORMPLAN
   -
   dbPlan=(
      -
         (name=prod, share=8, flashcache=on, flashlog=on),
         (name=dw, share=6, flashcache=on, flashlog=on),
         (name=prod_test, share=2, flashcache=off, flashlog=off),
         (name=prod_dev, share=1, flashcache=off, flashlog=off),
         (name=dw_test, share=2, flashcache=on, flashlog=off),
         (name=other, share=1))
```

You do not need to explicitly set `flashcache=on` or `flashlog=on` because they are the default settings for these directives.

### 6.4.9 Controlling Access to PMEM Cache and PMEM Log

You can use IORM to manage access to persistent memory (PMEM) cache and PMEM log for specific databases.

You can use the `pmemcache` and `pmemlog` directives in the interdatabase IORM plan to control access to PMEM cache and PMEM log.

The following example shows how to disable PMEM cache and PMEM log for the `prod_test` and `prod_dev` databases. In the example, PMEM log is also disabled for the `dw_test` database.

```
CellCLI> ALTER IORMPLAN
   -
   dbPlan=(
      -
         (name=prod, pmemcache=on, pmemlog=on),
         (name=dw, pmemcache=on, pmemlog=on),
         (name=prod_test, pmemcache=off, pmemlog=off),
         (name=prod_dev, pmemcache=off, pmemlog=off),
         (name=dw_test, pmemcache=on, pmemlog=off))
```
You can also use these directives in conjunction with other directives. For example:

```
CellCLI> ALTER IORMPLAN
-

dbPlan={
    (name=prod, share=8, pmemcache=on, pmemlog=on),
    (name=dw, share=6, pmemcache=on, pmemlog=on),
    (name=prod_test, share=2, pmemcache=off, pmemlog=off),
    (name=prod_dev, share=1, pmemcache=off, pmemlog=off),
    (name=dw_test, share=2, pmemcache=on, pmemlog=off),
    (name=other, share=1))
```

You do not need to explicitly set `pmemcache=on` or `pmemlog=on` because they are the default settings for these directives.

### 6.4.10 Listing an I/O Resource Management Plan

You can view the current interdatabase plan for a storage server using the CellCLI `LIST IORMPLAN` command on the storage server.

#### Example 6-17 Displaying Interdatabase Plan Details

This example shows how to get a detailed list of the interdatabase plan attributes.

```
CellCLI> LIST IORMPLAN DETAIL
    name:                  cell01_IORMPLAN
    status:                active
    catPlan:               name=administrative,level=1,allocation=80
                            name=interactive,level=2,allocation=90
                            name=batch,level=3,allocation=80
                            name=maintenance,level=4,allocation=50
                            name=other,level=4,allocation=50
    dbplan:                name=sales_prod, share=8, role=primary
                            name=sales_prod, share=1, limit=50,
                            role=standby
                            name=sales_test, share=1, limit=25
                            name=default, share=2
    objective:             balanced
```

**Related Topics**

- `LIST IORMPLAN`

### 6.4.11 Verifying the Configuration of I/O Resource Management

Use this checklist to verify that I/O Resource Management (IORM) is configured correctly.
• Verify that the following criteria are met when using IORM to manage I/O resources within a database:
  – A resource plan has been enabled.
  – The same resource plan has been enabled on all database instances.
    * If Oracle Database Resource Manager is enabled using Scheduler Window, then the same plan is always enabled on all database instances.
    * If Oracle Database Resource Manager is enabled using the RESOURCE_MANAGER_PLAN parameter, then use sid='*' to set the parameter for all database instances.
  – The resource plan includes MGMT_P[1-8] directives for each consumer group in the resource plan.

The following query can be used to verify the preceding criteria have been met:

```sql
SELECT DECODE(count(*), 0, 'Intra-Instance IORM Plan Enabled', 'No Intra-Instance IORM Plan Enabled') status
FROM gv$instance
WHERE inst_id NOT IN
  (SELECT inst_id FROM gv$rsrsrc_plan WHERE cpu_managed = 'ON');
```

• Verify that the interdatabase plan has been configured properly when IORM is used to manage I/O resources from multiple databases

```
CellCLI> LIST IORMPLAN DETAIL
```

If no interdatabase plan has been configured, then use the CellCLI ALTER IORMPLAN command to configure a plan. Each active database should have its own directive in the dbPlan parameter.

• Verify that sessions are mapped to the correct consumer group.

Run the following query while a workload is running:

```sql
SELECT r.sid,
  c.consumer_group current_consumer_group
FROM v$rsrsrc_session_info r, dba_rsrc_consumer_groups c
WHERE r.current_consumer_group_id = c.consumer_group_id
UNION
SELECT sid, 'OTHER_GROUPS' from v$rsrsrc_session_info
WHERE current_consumer_group_id = 0;
```

A session may not be in the expected consumer group due to the following configuration errors:

– **Missing privilege**: In order for a session to switch into a consumer group, its user or role must have permission to switch into that consumer group. The following query shows the permissions for all consumer groups.

```sql
SELECT grantee, granted_group
FROM DBA_RSRC_CONSUMER_GROUP_PRIVS
ORDER BY granted_group;
```
Use the following command to grant permission for any session to switch into the consumer group. This example shows how grant the permission for BATCH_GROUP.

```
EXEC dbms_resource_manager_privs.grant_switch_consumer_group -
     ('public', 'BATCH_GROUP', FALSE);
```

**Inactive consumer group:** If a session maps to or is manually switched to a consumer group that is not part of the current resource plan, then the session is switched into the default consumer group, OTHER_GROUPS.

If sessions are being assigned to consumer groups using mapping rules, then the following query can be used to determine the consumer group that the mapping rules selected, the mapping attribute that was used, and the consumer group that the session started in originally.

```
SELECT r.sid,
     r.mapped_consumer_group,
     r.mapping_attribute,
     c.consumer_group original_consumer_group
FROM v$rsrc_session_info r, dba_rsrc_consumer_groups c
WHERE r.orig_consumer_group_id = c.consumer_group_id;
```

If the mapped consumer group differs from the original consumer group, then the mapped consumer group was not part of the resource plan.

- Use CellCLI to list the number of small and large I/O requests that were issued for each consumer group across all databases.

```
CellCLI> LIST METRICCURRENT CG_IO_RQ_LG, CG_IO_RQ_SM   ATTRIBUTES
     name, -
     metricObjectName, metricValue, collectionTime;
```

Each consumer group that has an active I/O workload should generate small or large I/O requests according to these metrics.

- While a workload is running, verify that I/O loads are being managed in the correct consumer groups.

The following CellCLI command lists the number of small and large I/O requests that were issued for each consumer group across all databases:

```
CellCLI> LIST METRICCURRENT CG_IO_RQ_LG, CG_IO_RQ_SM   ATTRIBUTES
     name, -
     metricObjectName, metricValue, collectionTime;
```

- While the workload is running, query the actual I/O utilization for each category, database and consumer group.

The following CellCLI command lists the small and large I/O utilization for each database running on Oracle Exadata Storage Server:

```
CellCLI> LIST METRICCURRENT DB_IO_UTIL_LG, DB_IO_UTIL_SM ATTRIBUTES
     name, -
     metricObjectName, metricValue, collectionTime;
```
The output shows the percentage of disk resources utilized by small and large requests from the databases.

Related Topics

- **Monitoring IORM Utilization**
  You can use metrics to monitor IORM utilization.

### 6.4.12 Resetting Default Values in an Interdatabase Plan

To reset an interdatabase plan attribute to the default value, set the attribute to an empty string.

You can reset the entire plan, or separately reset the `catPlan` or `dbPlan`.

```
CellCLI> ALTER IORMPLAN dbPlan="", catPlan=""
CellCLI> ALTER IORMPLAN dbPlan=""
CellCLI> ALTER IORMPLAN catPlan=""
```
7 Monitoring Exadata

7.1 Introducing Exadata Monitoring Tools and Information Sources

You can monitor Exadata system performance by using various tools and information sources. The following topics introduce and compare the main Exadata performance monitoring tools and information sources:

- **Automatic Workload Repository (AWR)**
  AWR is the integrated performance diagnostic tool built into Oracle Database. AWR includes Exadata storage server configuration, health, and statistics information.

- **Database Statistics and Wait Events**
  Oracle Database provides performance information in a series of dynamic performance views (also known as $Vs$ views).

- **Exadata Metrics**
  Exadata metrics are recorded observations of important properties or values relating to the Exadata system software.

- **Exadata Alerts**
  Alerts draw attention to potential and actual problems and other events of interest to an administrator.

- **ExaWatcher**
  ExaWatcher collects very detailed system statistics, primarily focused on the operating system and network interfaces.

### 7.1.1 Automatic Workload Repository (AWR)

AWR is the integrated performance diagnostic tool built into Oracle Database. AWR includes Exadata storage server configuration, health, and statistics information.

AWR collects information from various database dynamic performance views at regular intervals. Each data collection is an AWR snapshot. An AWR report shows statistics captured between two AWR snapshots.

AWR collects and reports a wealth of statistical data related to Oracle Database, including database statistics and wait events from $V\text{SYSSTAT}$, $V\text{SYSTEM\_EVENT}$, $V\text{SEGMENT\_STATISTICS}$, $V\text{SQL}$. 
Commencing with Oracle Database release 12.1.0.2, AWR also collects performance statistics from Exadata storage servers and includes them in the AWR report. Consequently, AWR provides comprehensive and unified performance information that includes Oracle Database statistics and Exadata statistics, storage metrics, and alerts from all storage servers.

The AWR report contains the following Exadata-specific sections:

• Exadata Server Configuration
• Exadata Server Health Report
• Exadata Statistics

Key benefits include:

• AWR provides comprehensive and unified performance information that includes Oracle Database metrics and Exadata storage metrics and alerts from all storage servers.
• AWR enables simple outlier detection and analysis.
• The granularity of AWR information is customizable and based on the AWR collection interval, which is hourly by default.

It is important to note that the Exadata storage server statistics are collected and maintained by the storage servers, so the information is not restricted to a specific database or database instance. Consequently, the storage related statistics in AWR includes the I/O for all databases running on the storage servers, while the database statistics relate to the specific instance or database hosting AWR.

When using a container database (CDB), the Exadata-specific sections of the AWR report are only available in the root container.

As part of Oracle Database, AWR is constantly being enhanced with new statistics and analysis. The availability of a specific section or statistic is subject to the version of Oracle Database being used.

**Exadata Server Configuration**

The Exadata Server Configuration section of the AWR report shows configuration information from all of the associated storage servers, including model information, software versions, and storage information relating to cell disks, grid disks, ASM disk groups, and the IORM objective. By examining this information, you can easily see any configuration differences across the storage servers that could potentially affect the behavior or performance of the system.

The following example shows part of the Exadata Server Configuration section. The example is based on a system with 14 X8-2 High Capacity storage servers. In the example, all of the servers have the same software version. However, the storage information shows that one cell contains 2 flash cell disks, while the others all have 4 flash cell disks.
The Exadata Server Health Report section of the AWR report shows information about offline disks and open storage server alerts. When a disk is offline, then the total I/O bandwidth is reduced, which may affect I/O performance.

Alerts represent events of importance on a storage server that could affect its functionality and performance. Open alerts are alerts that have not been marked as examined by an administrator. For more details, see Exadata Alerts.

Ideally, there should be no open alerts in the AWR report, as shown in the following example:
Exadata Alerts Summary

No open alerts.

Back to Exadata Server Health Report

Exadata Alert Details

No open alerts.

Back to Exadata Server Health Report

The following example shows Exadata alert details indicating that the offload server was unable to start, which would affect smart scan performance on the system.

Exadata Statistics

AWR includes statistics from various Exadata storage server components, such as Smart I/O, Smart Flash Cache, Smart Flash Log, PMEM Cache, PMEM Log, and IO Resource Manager. It includes I/O statistics from both the operating system and the cell server software, and performs simple outlier analysis on these I/O statistics. Because Exadata typically distributes I/O operations evenly across all cells and all disks, outliers may indicate a potential problem and warrant further investigation. If any
cell or disk is behaving differently or performing more work, this could potentially cause slower performance for the whole system.

The Exadata Statistics section of the AWR report includes the following sub-sections:

- Performance Summary
- Exadata Resource Statistics
- Exadata Smart Statistics
  - Smart IO
  - Write Offload
  - Flash Log
  - Flash Cache
  - Memory Cache
  - PMEM Cache
- Exadata IO Reasons
  - Top IO Reasons by Requests
  - Top IO Reasons by MB
  - Internal IO Reasons
- Exadata Top Database Consumers
  - Top Databases by Requests
  - Top Databases by Requests - Details
  - Top Databases by Throughput
  - Top Databases by Requests per Cell
  - Top Databases by Requests per Cell - Details
  - Top Databases by Throughput per Cell
- Exadata Flash Wear

Key Exadata statistics are discussed in detail in the corresponding topics under Monitoring Oracle Exadata System Software Components.

Related Topics
- Automatic Workload Repository
- About Automatic Workload Repository Compare Periods Reports

### 7.1.2 Database Statistics and Wait Events

Oracle Database provides performance information in a series of dynamic performance views (also known as $V$ views).

The dynamic performance views contain a wealth of database statistics and wait events. By using dynamic performance views, a database administrator can perform
detailed and customized monitoring on individual SQL operations and database sessions. The dynamic performance views are also a major information source for Automatic Workload Repository (AWR) and other Oracle Database monitoring tools, such as SQL monitor.

The following sections introduce the main dynamic performance views that include Exadata-specific information.

- \texttt{V\$SYSSTAT}
- \texttt{V\$SESSION}
- \texttt{V\$SESSION\_EVENT}
- \texttt{V\$SYSTEM\_EVENT}
- \texttt{V\$ACTIVE\_SESSION\_HISTORY}
- \texttt{V\$SQL}
- \texttt{V\$SEGMENT\_STATISTICS}

\texttt{V\$SYSSTAT}

\texttt{V\$SYSSTAT} contains various system-wide statistics relating to the current database instance, including I/O statistics related to Exadata storage server.

For example, you can use the following query to display I/O statistics related to Exadata storage server:

\begin{verbatim}
SQL> SELECT name, value
        FROM v\$sysstat
        WHERE name like '%cell%' or name like 'physical'
        ORDER BY name;
\end{verbatim}

These statistics are also in the Global Activity Statistics or Instance Activity Statistics section of the AWR report.

\texttt{V\$SESSION}

\texttt{V\$SESSION} contains various statistics and wait events for currently active database sessions.

For example, you can use the following query to display a summary that shows the number of sessions that are currently waiting on cell events:

\begin{verbatim}
SQL> SELECT event, count(*)
        FROM v\$session
        WHERE event like 'cell'
            AND wait_time = 0  -- session currently waiting
        GROUP BY event
        ORDER BY count(*) desc;
\end{verbatim}

The following query displays event details for sessions that are currently waiting on cell events:

\begin{verbatim}
SQL> SELECT event, p1, p2, p3
        FROM v\$session
\end{verbatim}
WHERE event like 'cell%'
AND wait_time = 0;

For cell wait events in V$SESSION:

- The P1 column identifies the cell hash number. You can use the V$CELL view to identify the cell from the cell hash number.
- The P2 column identifies the disk hash number. You can use the V$ASM_DISK view to identify the disk name from the disk hash number.

V$SESSION_EVENT

V$SESSION_EVENT displays information on waits for an event by a session.

For example, you can use the following query to display a summary of the wait events for the current session. By examining the output from this query, you can compare the time spent waiting for Exadata storage server events with the time spent waiting for other events in the database.

```
SQL> SELECT event, total_waits, time_waited_micro, time_waited_micro/decode(total_waits, 0, null, total_waits) avg_wait_time
FROM v$session_event
WHERE wait_class != 'Idle'
AND sid = userenv('SID')
ORDER BY time_waited_micro desc;
```

V$SYSTEM_EVENT

V$SYSTEM_EVENT displays system-wide information on total waits for an event.

You can use the following query to display a summary of the wait events since start-up of the current database instance:

```
SQL> SELECT event, total_waits, time_waited_micro, time_waited_micro/decode(total_waits, 0, null, total_waits) avg_wait_time
FROM v$system_event
WHERE wait_class != 'Idle'
ORDER by time_waited_micro desc;
```

Information from V$SYSTEM_EVENT is contained in the Wait Events Statistics section of the AWR report.

V$ACTIVE_SESSION_HISTORY

V$ACTIVE_SESSION_HISTORY keeps a history of the active sessions on the system, where active is defined as using CPU or waiting on a non-idle wait event. For each session, V$ACTIVE_SESSION_HISTORY includes a wealth of information on what the session was doing, such as SQL execution plans and wait event details, including information about Exadata storage interactions.

For example, you can use the following query to display cell block reads over the past 5 minutes with latency in excess of 10ms:

```
SQL> SELECT sample_time, p1, p2, time_waited
```
FROM v$active_session_history
WHERE event = 'cell single block physical read'
    AND wait_time = 0  -- currently waiting
    AND sample_time > sysdate - 5/1440  -- past 5 minutes
    AND time_waited > 10000
ORDER by time_waited desc;

The Active Session History (ASH) Report section of the AWR report is based on
V$ACTIVE_SESSION_HISTORY. Oracle Database also provides a more detailed Active
Session History (ASH) report.

V$SQL, and related views like V$SQLAREA, V$SQLAREA_PLAN_HASH, V$SQLSTATS, and
V$SQLSTATS_PLAN_HASH, contain information and statistics for SQL statements
processed on the database instance.

V$SQL, and related views, include the following columns that contain information about
Exadata storage interactions:

- PHYSICAL_READ_BYTES — number of bytes read by the SQL
- PHYSICAL_READ_REQUESTS — number of read requests issued by the SQL
- PHYSICAL_WRITE_BYTES — number of bytes written by the SQL
- PHYSICAL_WRITE_REQUESTS — number of write requests issued by the SQL
- IO_CELL_OFFLOAD_ELIGIBLE_BYTES — number of bytes eligible for predicate
  offload to Exadata storage server
- IO_CELL_OFFLOAD_RETURNED_BYTES — number of bytes returned by smart scans
- IO_INTERCONNECT_BYTES — number of bytes exchanged between the database
  and the storage servers
- IO_CELL_UNCOMPRESSED_BYTES — number of uncompressed bytes read, where
  uncompressed bytes is the size after decompression
- OPTIMIZED_PHY_READ_REQUESTS — number of read requests satisfied from Exadata
  Smart Flash Cache, or read requests avoided due to storage index or columnar
  cache

The following example shows a query that displays Exadata offload processing
performance data. The output focuses on a query that scans the SALES table. The
output shows that all of the data in the SALES table (approximately 5 GB) was eligible
for offload processing. In this case, because of offload processing, even though the
query performed 5385.75 MB of I/O, only 417.65 MB of data was delivered over the
network to the database host.

```
SQL> SELECT sql_text,
             io_cell_offload_eligible_bytes/1024/1024 offload_eligible_mb,
             io_cell_uncompressed_bytes/1024/1024 io_uncompressed_mb,
             io_interconnect_bytes/1024/1024 io_interconnect_mb,
             io_cell_offload_returned_bytes/1024/1024 cell_return_bytes_mb,
             (physical_read_bytes + physical_write_bytes)/1024/1024
             io_disk_mb
      FROM v$sql
WHERE UPPER(sql_text) LIKE 'FROM SALES%';
```
Information from V$SQL is displayed in the SQL Statistics section of the AWR report.

V$SEGMENT_STATISTICS contains statistics on a per-segment basis. The segment-level statistics can be used to detect specific objects, such as tables or indexes, that are performing optimized reads from Exadata storage.

The optimized physical read segment statistic records the number of read requests for objects that are read from Exadata Smart Flash Cache or reads that are avoided through the use of storage index or columnar cache. The optimized physical writes statistic records the number of write requests for an object that first went to Exadata Smart Flash Cache. Such write requests can be synchronized to the disk later in a lazy manner to free up cache space.

The following example query displays objects that are associated with more than 1000 optimized reads from Exadata storage. A similar query can be used to determine the objects that are associated with very few optimized reads.

SQL> SELECT object_name, value
FROM v$segment_statistics
WHERE statistic_name='optimized physical reads'
AND value>1000
ORDER BY value;

Information from V$SEGMENT_STATISTICS is displayed in the Segment Statistics section of the AWR report.

### 7.1.3 Exadata Metrics

Exadata metrics are recorded observations of important properties or values relating to the Exadata system software.

Most metrics relate to the storage server and its components, such as flash cache, cell disks, and grid disks. Storage server metrics enable detailed monitoring of Exadata storage server performance.

Metrics observations are initially stored in memory and later stored in a disk-based repository for a more permanent history. Metric values typically record the absolute value of a cumulative counter or a rate of change in a cumulative counter over time. Some metrics are used to record the time of state transitions.

Metrics are of the following types:

- Cumulative metrics are statistics that accumulate over time since the metric was created or the server was restarted.
• Instantaneous metrics contain the current value at the time of the metric observation.

• Rate metrics are computed statistics where the value is averaged over time. Some metrics differentiate between small I/O and large I/O. For such metrics, small I/O means I/O that is less than or equal to 128 KB in size. Large I/O is greater than 128 KB in size.

You can use the CellCLI LIST command to display and monitor Exadata metrics. See Displaying Metrics.

Exadata metrics contain detailed statistics for most Exadata components. However, there is no inherent summary or report for Exadata metrics. Also, because Exadata metrics are managed within each Exadata server, metric observations must be collected and correlated by the Exadata administrator to gain a system-wide view.

**Related Topics**

- Displaying Metrics
  You can use the CellCLI LIST command to display and monitor metrics for cell objects.

### 7.1.4 Exadata Alerts

Alerts draw attention to potential and actual problems and other events of interest to an administrator.

There are three types of alerts; informational alerts, warning alerts, and critical alerts. Metrics observations are automatically compared with stored thresholds. When a metric crosses a specified threshold, an alert is automatically generated. Metrics can be associated with warning and critical thresholds. When a metric value crosses a warning threshold, then a warning alert is generated. When a metric crosses a critical threshold, then a critical alert is generated.

Exadata administrators can specify the warning and critical threshold for most metrics. There are also some preset thresholds. For example, there are built-in thresholds for ambient temperature. If the temperature is too low or too high, an alert is automatically generated.

Some system conditions and state changes can generate informational alerts; for example, a server chassis is opened. More serious system errors and state changes can also generate warning or critical alerts; for example, a hard disk device is reporting disk errors.

This system of thresholds and alerts enables administrators to focus on the most important events, while maintaining awareness other interesting events.

You can use CellCLI commands to manage and display Exadata alerts and thresholds. You can also configure alert notification using email and SNMP. Alert information is also included in the AWR report.

Like Exadata metrics, alerts are managed separately on each server.

**Related Topics**

- Monitoring Alerts
  You can monitor and receive notifications for alerts.
7.1.5 ExaWatcher

ExaWatcher collects very detailed system statistics, primarily focused on the operating system and network interfaces.

ExaWatcher runs separately on each server and collects server and network-related data using operating system tools and utilities, such as \texttt{iostat}, \texttt{vmstat}, \texttt{mpstat}, \texttt{top}. It also collects storage server statistics using \texttt{cellsrvstat}.

The information is very granular, with most operating system statistics collected at 5-second intervals. To extract statistics from ExaWatcher for analysis, you can use the \texttt{GetExaWatcherResults.sh} script. The output from \texttt{GetExaWatcherResults.sh} also include charts for a small subset of key statistics.

ExaWatcher is typically used for detailed debugging. Consequently, it is recommended to keep baselines of ExaWatcher data for comparison purposes.

Related Topics

- System Diagnostics Data Gathering with sosreports and Oracle ExaWatcher
  You can use the sosreport utility and Oracle ExaWatcher to diagnose problems with your system.

- About ExaWatcher Charts

7.2 Guidelines for Exadata Monitoring

Regardless of the information source, it is considered good practice to maintain baseline collections of the various performance data sources covering the important periods of your workload, such as peak periods, month-end or year-end processing, or before patching or upgrades.

Baseline information enables a comparison with a known 'normal' state when a performance regression occurs. Such comparison helps to identify changes and guide further analysis.

Baseline data should include AWR data (using AWR exports or AWR baselines), Exadata metrics, and ExaWatcher collections. For data warehousing workloads that rely on smart scans, keeping a collection of SQL monitor reports for key SQL statements is also highly recommended.

7.3 Monitoring Oracle Exadata System Software Components

- Monitoring Exadata Smart Flash Cache
- Monitoring PMEM Cache
- Monitoring Exadata Smart Flash Log
- Monitoring PMEM Log
- Monitoring Smart I/O
- Monitoring I/O Resource Management (IORM)
• Monitoring Cell Disk I/O
• Monitoring Grid Disks
• Monitoring Host Interconnect Metrics
  Host interconnect metrics provide information about the I/O transmission for hosts
  that access cell storage.
• Monitoring RAM Cache

7.3.1 Monitoring Exadata Smart Flash Cache

Exadata Smart Flash Cache holds frequently accessed data in flash storage, while
most data is kept in very cost-effective disk storage. Caching occurs automatically and
requires no user or administrator effort. Exadata Smart Flash Cache can intelligently
determine the data that is most useful to cache based on data usage, access patterns,
and hints from the database that indicate the type of data being accessed.

Exadata Smart Flash Cache can operate in Write-Through mode or Write-Back mode.
In Write-Through mode, database writes go to disk first and subsequently populate
Flash Cache. If a flash device fails with Exadata Smart Flash Cache operating in
Write-Through mode, there is no data loss because the data is already on disk.

In Write-Back mode, database writes go to Flash Cache first and later to disk. The
contents of the Write-Back Flash Cache is persisted across reboots, eliminating any
warm-up time needed to populate the cache. Write-intensive applications can benefit
from Write-Back caching by taking advantage of the fast latencies provided by flash.
The amount of disk I/O also reduces when the cache absorbs multiple writes to
the same block before writing it to disk. However, if a flash device fails while using
Write-Back mode, data that is not yet persistent to disk is lost and must be recovered
from a mirror copy. For this reason, Write-Back mode is recommended in conjunction
with high redundancy ASM disk groups.

Exadata Smart Flash Cache accelerates OLTP performance by providing fast I/O
for frequently accessed data blocks. It also accelerates Decision Support System
(DSS) performance by automatically caching frequently scanned data and temporary
segments, providing columnar cache storage, and enabling other features that
optimize the performance of large analytic queries and large loads.

On Extreme Flash (EF) storage servers, all of the data resides in flash. Consequently,
Exadata Smart Flash Cache is not required for normal caching. However, in this case
Exadata Smart Flash Cache is still used to host the Columnar Cache, which caches
data in columnar format and optimizes various analytical queries.

Performance issues related to Exadata Smart Flash Cache typically exhibit increased
cell single block physical read latencies in the database. However, occasional
long latencies associated with cell single block physical read do not necessarily
indicate a performance issue with Exadata Smart Flash Cache, and may simply
indicate that the read request was not satisfied using Exadata Smart Flash Cache.

• Monitoring Exadata Smart Flash Cache Using AWR
• Monitoring Exadata Smart Flash Cache Using Database Statistics and Wait
  Events
• Monitoring Exadata Smart Flash Cache Using Exadata Metrics
• What to Look For When Monitoring Exadata Smart Flash Cache
7.3.1.1 Monitoring Exadata Smart Flash Cache Using AWR

Automatic Workload Repository (AWR) contains a wealth of information relating to Exadata Smart Flash Cache. Following are descriptions and examples of key sections in the AWR report that contain information about Exadata Smart Flash Cache. By reviewing these sections of the AWR report, administrators can understand how Exadata Smart Flash Cache operates.

Flash Cache Configuration and Space Usage

The Flash Cache Configuration section contains summary information including the caching mode (Write-Through or Write-Back), status and overall size. The Flash Cache Space Usage section provides summary statistics on space usage in Exadata Smart Flash Cache.

The following example shows Exadata Smart Flash Cache configured in Write-Back mode on all cells, with a total size of almost 24 TB.

The space usage shows that approximately 3.9 TB of Exadata Smart Flash Cache space is in use on each cell. Nearly all of the allocated space is being used for OLTP, which is typically data that is read in block mode into the database buffer cache. There is also a small amount of space allocated to large writes and scan workloads.

---

**Figure 7-4** AWR Report: Flash Cache Configuration and Space Usage

### Flash Cache Configuration
- These statistics are collected by the cells and are not restricted to this database or instance
- Size (GB) - configured size for Flash Cache

<table>
<thead>
<tr>
<th>Mode</th>
<th>Compression</th>
<th>Status</th>
<th>Size (GB)</th>
<th>Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>WriteBack</td>
<td>normal</td>
<td></td>
<td>23945.81</td>
<td>All</td>
</tr>
</tbody>
</table>

Back to Exadata Flash Cache
Back to Exadata Smart Statistics

### Flash Cache Space Usage
- These statistics are collected by the cells and are not restricted to this database or instance
- Space is at the time of the end snapshot
- Ordered by Space (GB) desc

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>Space (GB)</th>
<th>OLTP</th>
<th>%Clean</th>
<th>%Unflushed</th>
<th>%Temp Spill</th>
<th>%Write Only</th>
<th>%Clean</th>
<th>%Columnar</th>
<th>OLTP</th>
<th>%Scan</th>
<th>%Unflushed</th>
<th>%Temp Spill</th>
<th>%Write Only</th>
<th>%Clean</th>
<th>%Columnar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (3)</td>
<td>11,732.92</td>
<td></td>
<td>4.84</td>
<td>93.34</td>
<td>1.11</td>
<td>0.46</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm/cell10</td>
<td>3,948.12</td>
<td></td>
<td>5.97</td>
<td>92.71</td>
<td>0.85</td>
<td>0.21</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm/cell11</td>
<td>3,918.71</td>
<td></td>
<td>5.00</td>
<td>92.26</td>
<td>1.07</td>
<td>0.92</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm/cell12</td>
<td>3,866.08</td>
<td></td>
<td>3.51</td>
<td>95.07</td>
<td>0.91</td>
<td>0.25</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Back to Exadata Flash Cache
Back to Exadata Smart Statistics

---

**Flash Cache User Reads**

The Flash Cache User Reads sections show information about read requests, read throughput, and read efficiency from database clients. The statistics show the I/Os against different areas of Exadata Smart Flash Cache:

- OLTP - relates to block requests
- Scan - relates to scan requests
Columnar - relates to columnar cache requests

Keep - relates to read requests against the KEEP pool

Figure 7-5  AWR Report: Flash Cache User Reads

Flash Cache User Reads

- These statistics are collected by the cells and are not restricted to this database or instance
- Total - total number of requests from Flash Cache
- OLTP/Scan/Columnar reads include reads on keep objects
- Misses/Partial Hits - number of read request misses and partial hits
- Ordered by Total Hit Read Requests desc:

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>Total Hits</th>
<th>Hits per Sec</th>
<th>OLTP</th>
<th>Scan</th>
<th>Columnar</th>
<th>Keep</th>
<th>Misses</th>
<th>Partial Hits</th>
<th>Active Secondary Hits</th>
<th>Active Secondary Misses</th>
<th>Total Hits</th>
<th>Hit Misses</th>
<th>OLTP</th>
<th>Scan</th>
<th>Columnar</th>
<th>Keep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>184,127</td>
<td>39.85</td>
<td>184,020</td>
<td>107</td>
<td>28.331</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11,506.77</td>
<td>5.86</td>
<td>11,486.36</td>
<td>20.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db0001cell01</td>
<td>79.162</td>
<td>40.30</td>
<td>79.134</td>
<td>38</td>
<td>6.423</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.987.43</td>
<td>2.52</td>
<td>4.987.37</td>
<td>7.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db0001cell10</td>
<td>62.978</td>
<td>27.00</td>
<td>62.960</td>
<td>26</td>
<td>7.320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.309.45</td>
<td>1.69</td>
<td>3.304.27</td>
<td>6.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db0001cell12</td>
<td>51.946</td>
<td>29.50</td>
<td>51.946</td>
<td>43</td>
<td>12.488</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.249.89</td>
<td>1.96</td>
<td>3.241.82</td>
<td>6.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Flash Cache User Reads Per Second

- These statistics are collected by the cells and are not restricted to this database or instance
- Total - total number of reads per second from Flash Cache
- OLTP/Scan/Columnar reads include reads on keep objects
- Ordered by Total Hit Read Requests per Second desc:

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>OLTP</th>
<th>Scan</th>
<th>Columnar</th>
<th>Keep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>13.37</td>
<td>5.82</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>db0001cell01</td>
<td>40.30</td>
<td>0.02</td>
<td>3.27</td>
<td>2.52</td>
</tr>
<tr>
<td>db0001cell10</td>
<td>27.00</td>
<td>0.01</td>
<td>3.73</td>
<td>1.65</td>
</tr>
<tr>
<td>db0001cell12</td>
<td>26.98</td>
<td>0.48</td>
<td>6.36</td>
<td>1.61</td>
</tr>
</tbody>
</table>

Back to Exadata Flash Cache
Back to Exadata Smart Statistics

Flash Cache User Reads Efficiency

- These statistics are collected by the cells and are not restricted to this database or instance
- Ordered by Total Hit Reads desc:

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>OLTP</th>
<th>Scan</th>
<th>Columnar</th>
<th>Keep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>184,127</td>
<td>11,506.77</td>
<td>184,020</td>
<td>5.86</td>
</tr>
<tr>
<td>db0001cell01</td>
<td>79.162</td>
<td>4.987.43</td>
<td>79.134</td>
<td>2.52</td>
</tr>
<tr>
<td>db0001cell10</td>
<td>62.978</td>
<td>3.309.45</td>
<td>62.960</td>
<td>1.69</td>
</tr>
<tr>
<td>db0001cell12</td>
<td>51.946</td>
<td>3.249.89</td>
<td>51.946</td>
<td>1.96</td>
</tr>
</tbody>
</table>

Back to Exadata Flash Cache
Back to Exadata Smart Statistics

Flash Cache User Writes

The Flash Cache User Writes section shows information about write requests and write throughput for Exadata Smart Flash Cache in Write-Back mode.

In this section, First Writes indicate new data being written to Exadata Smart Flash Cache, while Overwrites indicate data being overwritten in Exadata Smart Flash Cache. First Writes also require writing additional flash cache metadata. Overwrites represents the disk writes that were avoided by using Exadata Smart Flash Cache in Write-Back mode.

Figure 7-6  AWR Report: Flash Cache User Writes

Flash Cache User Writes

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>Total Hits</th>
<th>OLTP</th>
<th>Scan</th>
<th>Columnar</th>
<th>Keep</th>
<th>Misses</th>
<th>Total Hits</th>
<th>Hit Misses</th>
<th>OLTP</th>
<th>Scan</th>
<th>Columnar</th>
<th>Keep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>184,127</td>
<td>11,506.77</td>
<td>184,020</td>
<td>5.86</td>
<td>11,486.36</td>
<td>20.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>db0001cell01</td>
<td>79.162</td>
<td>4.987.43</td>
<td>79.134</td>
<td>2.52</td>
<td>4.987.37</td>
<td>7.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>db0001cell10</td>
<td>62.978</td>
<td>3.309.45</td>
<td>62.960</td>
<td>1.69</td>
<td>3.304.27</td>
<td>6.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>db0001cell12</td>
<td>51.946</td>
<td>3.249.89</td>
<td>51.946</td>
<td>1.96</td>
<td>3.241.82</td>
<td>6.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Back to Exadata Flash Cache
Back to Exadata Smart Statistics
Flash Cache User Writes - Large Writes

The Flash Cache User Writes - Large Writes sections show information about large write requests that are absorbed and rejected by Exadata Smart Flash Cache in Write-Back mode.

Figure 7-7  AWR Report: Flash Cache User Writes - Large Writes

Flash Cache Internal Reads

The Flash Cache Internal Reads section shows reads from Exadata Smart Flash Cache that are performed by Oracle Exadata System Software. These statistics are populated when Exadata Smart Flash Cache is in Write-Back mode.

The Disk Writer IO Detail columns relate to internal I/Os performed to persist data from Exadata Smart Flash Cache to hard disk devices. These columns show both the reads from flash and the writes to hard disk.

Figure 7-8  AWR Report: Flash Cache Internal Reads

Flash Cache Internal Reads

- These statistics are collected by the cells and are not restricted to this database or instance
- Read to Disk Write - reads from flash cache to write to hard disk
- Disk Writer IO Detail - actual number of I/Os
- Ordered by Total Read Req desc

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>Total</th>
<th>Total</th>
<th>Reads from Flash</th>
<th>Requests/s</th>
<th>MB/s</th>
<th>Writes to Hard Disk</th>
<th>Requests/s</th>
<th>MB/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (3)</td>
<td>508,857</td>
<td>259.36</td>
<td>463,909.70</td>
<td>236.45</td>
<td>3,795.31</td>
<td>237.21</td>
<td>394.90</td>
<td>236.47</td>
</tr>
<tr>
<td>dbm0celadm11</td>
<td>170,670</td>
<td>86.99</td>
<td>155,562.63</td>
<td>79.29</td>
<td>1,272.40</td>
<td>79.53</td>
<td>131.43</td>
<td>79.29</td>
</tr>
<tr>
<td>dbm0celadm12</td>
<td>170,668</td>
<td>86.99</td>
<td>155,198.30</td>
<td>79.10</td>
<td>1,270.25</td>
<td>79.39</td>
<td>133.90</td>
<td>79.11</td>
</tr>
<tr>
<td>dbm0celadm10</td>
<td>168,119</td>
<td>85.69</td>
<td>153,148.77</td>
<td>78.06</td>
<td>1,252.67</td>
<td>78.29</td>
<td>129.57</td>
<td>78.07</td>
</tr>
</tbody>
</table>
The Flash Cache Internal Writes section shows writes to Exadata Smart Flash Cache that are performed by Oracle Exadata System Software. The internal writes are I/Os that populate Exadata Smart Flash Cache in response to a cache read miss.

The statistics also include metadata writes that occur when Exadata Smart Flash Cache is in Write-Back mode. Metadata writes occur when a cacheline is used to cache new data.

7.3.1.2 Monitoring Exadata Smart Flash Cache Using Database Statistics and Wait Events

The following table describes various database statistics that are useful for monitoring Exadata Smart Flash Cache. The statistics are available in various dynamic performance views, including `V$SYSSTAT`, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell flash cache read hits</td>
<td>The number of read requests that were satisfied by the cache</td>
</tr>
<tr>
<td>cell flash cache read hits for smart IO</td>
<td>The number of read requests for smart IO that were satisfied by the cache</td>
</tr>
<tr>
<td>cell flash cache read hits for temp IO</td>
<td>The number of read requests for temp IO that were satisfied by the cache</td>
</tr>
<tr>
<td>cell flash read hits for controlfile reads</td>
<td>The number of read requests for a database control file that were satisfied by the cache</td>
</tr>
<tr>
<td>cell overwrites in flash</td>
<td>Total number of write requests that overwrote an existing cacheline in Exadata Smart Flash Cache that had not been written to disk. In effect, this is the amount of disk I/O saved by using Write-Back mode. This statistic is incremented once per mirror write.</td>
</tr>
<tr>
<td>cell partial writes in flash</td>
<td>Total number of write requests written to both Exadata Smart Flash Cache and disk. Part of the data was written to flash, and the rest was written to disk. This statistic is incremented once per mirror write.</td>
</tr>
<tr>
<td>Statistic</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>cell writes to flash cache</td>
<td>Total number of write requests written entirely to Exadata Smart Flash Cache. This statistic is incremented once per mirror write.</td>
</tr>
<tr>
<td>cell writes to flash cache for temp IO</td>
<td>The number of write requests for temporary segments that were absorbed by Exadata Smart Flash Cache.</td>
</tr>
<tr>
<td>physical read IO requests</td>
<td>Number of physical read requests issued by the database.</td>
</tr>
<tr>
<td>physical read requests optimized</td>
<td>Total number of read requests satisfied by using Exadata Smart Flash Cache, and read requests avoided by using storage index or columnar cache.</td>
</tr>
<tr>
<td>physical read total bytes</td>
<td>Total amount of I/O bytes for reads issued by the database, whether or not the read was offloaded to the storage servers.</td>
</tr>
<tr>
<td>physical read total bytes optimized</td>
<td>Total number of bytes read from Exadata Smart Flash Cache, and bytes avoided by using storage index or columnar cache.</td>
</tr>
<tr>
<td>physical read total IO requests</td>
<td>Total number of read requests issued by the database for all activity including application, backup, recovery, and other utilities.</td>
</tr>
<tr>
<td>physical write IO requests</td>
<td>Number of physical write requests issued by the database.</td>
</tr>
<tr>
<td>physical write total bytes</td>
<td>Total number of IO bytes for writes issued by the database for all activity.</td>
</tr>
<tr>
<td>physical write total bytes optimized</td>
<td>Total number of bytes written to Exadata Smart Flash Cache. These writes are synchronized to disk in a lazy manner.</td>
</tr>
<tr>
<td>physical write total IO requests</td>
<td>Total number of write requests issued by the database for all activity, including application, backup, recovery and other utilities.</td>
</tr>
</tbody>
</table>

The following table describes database wait events that are useful for monitoring Exadata Smart Flash Cache. The wait events are visible in various dynamic performance views, including V$SESSION, V$SYSTEM_EVENT and V$SESSION_EVENT, and may be displayed in the Wait Event sections of the AWR report.

The latency of the cell single block physical read event typically indicates if the read is satisfied from Exadata Smart Flash Cache. The higher the latency, the higher the likelihood that the request was not satisfied from Exadata Smart Flash Cache.
<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell list of blocks physical read</td>
<td>This wait event occurs during recovery or during buffer pre-fetching (rather than performing multiple single-block reads). It is used to monitor database blocks that must be changed as part of recovery and are read in parallel for the database. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the number of blocks processed during the I/O read operation. This wait event is equivalent to db file parallel read for a cell.</td>
</tr>
<tr>
<td>cell multiblock physical read</td>
<td>The wait time is the actual time it takes to do all of the I/Os when reading multiple data blocks. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the total number of bytes processed during the I/O read operation. This wait event is equivalent to db file scattered read for a cell.</td>
</tr>
<tr>
<td>cell single block physical read</td>
<td>The wait time is the actual time it takes to do the I/O for a single block read. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the number of bytes processed during the I/O read operation. This wait event is equivalent to db file sequential read for a cell.</td>
</tr>
<tr>
<td>cell interconnect retransmit during physical read</td>
<td>This wait event appears during retransmission for an I/O of a single-block or multiblock read. The cell hash number in the P1 column in the V$SESSION_WAIT view is the same cell identified for cell single block physical read and cell multiblock physical read. The P2 column contains the subnet number to the cell, and the P3 column contains the number of bytes processed during the I/O read operation.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.

7.3.1.3 Monitoring Exadata Smart Flash Cache Using Exadata Metrics

Exadata metrics that are related to Exadata Smart Flash Cache are identified in the Exadata storage server METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=FLASHCACHE.
Example 7-1  Displaying Flash Cache Metric Definitions

This example shows how to display the flash cache metric definitions that are available in the Oracle Exadata System Software.

```
CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE
OBJECTTYPE = FLASHCACHE
   FC_BYKEEP_DIRTY                         "Number of megabytes unflushed for keep objects on FlashCache"
   FC_BYKEEP_OLTP                          "Number of megabytes for OLTP keep objects in flash cache"
   FC_BYKEEP_OVERWR                        "Number of megabytes pushed out of the FlashCache because of space limit for keep objects"
   FC_BYKEEP_OVERWR_SEC                    "Number of megabytes per second pushed out of the FlashCache because of space limit for keep objects"
   FC_BYKEEP_USED                         "Number of megabytes used for keep objects on FlashCache"
   FC_BY_ALLOCATED                        "Number of megabytes allocated in flash cache"
   FC_BY_ALLOCATED_DIRTY                   "Number of megabytes allocated for unflushed data in flash cache"
   FC_BY_ALLOCATED_OLTP                   "Number of megabytes allocated for OLTP data in flash cache"
   FC_BY_DIRTY                            "Number of unflushed megabytes in FlashCache"
   FC_BY_STALE_DIRTY                       "Number of unflushed megabytes in FlashCache which cannot be flushed because cached disks are not accessible"
   FC_BY_USED                             "Number of megabytes used on FlashCache"
   FC_COL_BYKEEP_USED                     "Number of megabytes used for keep objects in Columnar FlashCache"
   FC_COL_BY_USED                         "Number of megabytes used in Columnar FlashCache"
   FC_COL_IO_BYKEEP_R                     "Number of megabytes read from Columnar FlashCache for keep objects"
   FC_COL_IO_BYKEEP_R_SEC                 "Number of megabytes per second read from Columnar FlashCache for keep objects"
   FC_COL_IO_BY_R                         "Number of megabytes that were read from Columnar FlashCache"
   FC_COL_IO_BY_R_ELIGIBLE                "Number of megabytes eligible to read from Columnar FlashCache"
   FC_COL_IO_BY_R_ELIGIBLE_SEC           "Number of megabytes per second eligible to read from Columnar FlashCache"
   FC_COL_IO_BY_R_SEC                     "Number of megabytes per second that were read from Columnar FlashCache"
   FC_COL_IO_BY_SAVED                     "Number of megabytes saved by reads from Columnar FlashCache"
   FC_COL_IO_BY_SAVED_SEC                 "Number of megabytes saved per second by reads from Columnar FlashCache"
   FC_COL_IO_BY_W_POPULATE                "Number of megabytes that are population writes into Columnar FlashCache due to read miss"
   FC_COL_IO_BY_W_POPULATE_SEC            "Number of megabytes per second that are population writes into Columnar FlashCache due to
```
read miss
FC_COL_IO_RQKEEP_R  "Number of requests read for keep objects from Columnar FlashCache"
FC_COL_IO_RQKEEP_R_SEC  "Number of requests read per second for keep objects from Columnar FlashCache"
FC_COL_IO_RR  "Number of requests that were read from Columnar FlashCache"
FC_COL_IO_RR_ELIGIBLE  "Number of reads eligible for Columnar FlashCache"
FC_COL_IO_RR_ELIGIBLE_SEC  "Number of reads per second that were read from Columnar FlashCache"
FC_COL_IO_RR_SEC  "Number of requests per second that were read from Columnar FlashCache"
FC_COL_IO_RW_POPULATE  "Number of requests that are population writes into Columnar FlashCache due to read miss"
FC_COL_IO_RW_POPULATE_SEC  "Number of requests per second that are population writes into Columnar FlashCache due to read miss"

read from FlashCache for keep objects
FC_IO_BYKEEP_R  "Number of megabytes"
FC_IO_BYKEEP_R_SEC  "Number of megabytes per second"

written to FlashCache for keep objects
FC_IO_BYKEEP_W  "Number of megabytes"
FC_IO_BYKEEP_W_SEC  "Number of megabytes per second"

written from flash cache to hard disks
FC_IO_BY_DISK_WRITE  "Number of megabytes"
FC_IO_BY_DISK_WRITE_SEC  "Number of megabytes per second"

of small reads (OLTP) from flash cache
FC_IO_BY_R_ACTIVE_SECONDARY  "Number of megabytes for active secondary reads satisfied from flash cache"
FC_IO_BY_R_ACTIVE_SECONDARY_MISS  "Number of megabytes for active secondary reads not satisfied from flash cache"
FC_IO_BY_R_ACTIVE_SECONDARY_MISS_SEC  "Number of megabytes per second for active secondary reads not satisfied from flash cache"
FC_IO_BY_R_DISK_WRITER  "Number of megabytes read from flash cache by disk writer"
FC_IO_BY_R_DISK_WRITER_SEC  "Number of megabytes per second read from flash cache by disk writer"
FC_IO_BY_R_RW  "Number of megabytes of large reads (DW) from flash cache"
FC_IO_BY_R_RW_SEC  "Number of megabytes of large reads (DW) per second from flash cache"
FC_IO_BY_R_MISS  "Number of megabytes of small reads (OLTP) from disks because some of the requested data was not in flash cache"
FC_IO_BY_R_MISS_RW  "Number of megabytes of large reads (DW) from disks because some of the requested data was not in flash cache"
of large reads (DW) per second from disks because some of the requested data was not in flash cache"

FC_IO_BY_R_MISS_SEC "Number of megabytes of small reads (OLTP) per second from disks because some of the requested data was not in flash cache"

FC_IO_BY_R_SEC "Number of megabytes of small reads (OLTP) per second from flash cache"

FC_IO_BY_R_SKIP "Number of megabytes read from disks for IO requests that bypass FlashCache"

FC_IO_BY_R_SKIP_FC_THROTTLE "Number of megabytes read from disk for IO requests that bypass FlashCache due to heavy load on FlashCache"

FC_IO_BY_R_SKIP_FC_THROTTLE_SEC "Number of megabytes read per second from disk for IO requests that bypass FlashCache due to heavy load on FlashCache"

FC_IO_BY_R_SKIP_LG "Number of megabytes read from disk for IO requests that bypass FlashCache due to the large IO size"

FC_IO_BY_R_SKIP_LG_SEC "Number of megabytes read per second from disk for IO requests that bypass FlashCache due to the large IO size"

FC_IO_BY_R_SKIP_NCMIRROR "Number of megabytes read from disks per second for IO requests that bypass FlashCache as the IO is on non-primary, non-active secondary mirror"

FC_IO_BY_R_SKIP_SEC "Number of megabytes read from disks per second for IO requests that bypass FlashCache"

FC_IO_BY_W "Number of megabytes written to FlashCache"

FC_IO_BY_W_DISK_WRITER "Number of megabytes written to hard disks by disk writer"

FC_IO_BY_W_DISK_WRITER_SEC "Number of megabytes per second written to hard disks by disk writer"

FC_IO_BY_W_FIRST "Number of megabytes that are first writes into flash cache"

FC_IO_BY_W_FIRST_SEC "Number of megabytes per second that are first writes into flash cache"

FC_IO_BY_W_LG_CHINT "Number of megabytes used for large writes to flash cache due to cache hint"

FC_IO_BY_W_LG_DTAGE "Number of megabytes used for data aging large writes to flash cache"

FC_IO_BY_W_LG_MRCV "Number of megabytes used for media recovery large writes to flash cache"

FC_IO_BY_W_METADATA "Number of megabytes that are flash cache metadata writes"

FC_IO_BY_W_METADATA_SEC "Number of megabytes per second that are flash cache metadata writes"

FC_IO_BY_W_OVERWRITE "Number of megabytes that are overwrites into flash cache"

FC_IO_BY_W_OVERWRITE_SEC "Number of megabytes per second that are overwrites into flash cache"

FC_IO_BY_W_POPULATE "Number of megabytes that are population writes into flash cache due to read miss"

FC_IO_BY_W_POPULATE_SEC "Number of megabytes per second that are population writes into flash cache due to read miss"
<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC_IO_BY_W_SKIP</td>
<td>Number of megabytes written to disk for IO requests that bypass FlashCache</td>
</tr>
<tr>
<td>FC_IO_BY_W_SKIP_FC_THROTTLE</td>
<td>Number of megabytes written to disk for IO requests that bypass FlashCache due to heavy load on FlashCache</td>
</tr>
<tr>
<td>FC_IO_BY_W_SKIP_FC_THROTTLE_SEC</td>
<td>Number of megabytes written per second to disk for IO requests that bypass FlashCache due to heavy load on FlashCache</td>
</tr>
<tr>
<td>FC_IO_BY_W_SKIP_LG</td>
<td>Number of megabytes written to disk for IO requests that bypass FlashCache due to the large IO size</td>
</tr>
<tr>
<td>FC_IO_BY_W_SKIP_LG_SEC</td>
<td>Number of megabytes written per second to disk for IO requests that bypass FlashCache due to the large IO size</td>
</tr>
<tr>
<td>FC_IO_BY_W_SKIP_NCMIRROR</td>
<td>Number of megabytes written to disk for IO requests that bypass FlashCache as the IO is on non-primary, non-active secondary mirror</td>
</tr>
<tr>
<td>FC_IO_BY_W_SKIP_SEC</td>
<td>Number of megabytes written to disk per second for IO requests that bypass FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_R</td>
<td>Number of read requests for keep objects from FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_R_MISS</td>
<td>Number of read requests for keep objects which did not find all data in FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_R_MISS_SEC</td>
<td>Number of read requests per second for keep objects which did not find all data in FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_R_SKIP</td>
<td>Number of read requests for keep objects that bypass FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_R_SKIP_SEC</td>
<td>Number of read requests per second for keep objects that bypass FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_R_SEC</td>
<td>Number of read requests per second for keep objects from FlashCache</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_W</td>
<td>Number of requests for keep objects which resulted in FlashCache being populated with data</td>
</tr>
<tr>
<td>FC_IO_RQKEEP_W_SEC</td>
<td>Number of requests per second for keep objects which resulted in FlashCache being populated with data</td>
</tr>
<tr>
<td>FC_IO_RQ_DISK_WRITE</td>
<td>Number of requests written from flash cache to hard disks</td>
</tr>
<tr>
<td>FC_IO_RQ_DISK_WRITE_SEC</td>
<td>Number of requests per second written from flash cache to hard disks</td>
</tr>
<tr>
<td>FC_IO_RQ_R</td>
<td>Number of small reads (OLTP) satisfied from the flash cache</td>
</tr>
<tr>
<td>FC_IO_RQ_REPLACEMENT_ATTEMPTED</td>
<td>Number of requests attempted to find space in the flash cache</td>
</tr>
<tr>
<td>FC_IO_RQ_REPLACEMENT_DW_FAILED</td>
<td>Number of times that client DW IOs failed to find a replacement buffer</td>
</tr>
<tr>
<td>FC_IO_RQ_REPLACEMENT_DW_SUCCEEDED</td>
<td>Number of times that client DW IOs succeeded in finding a replacement buffer</td>
</tr>
<tr>
<td>FC_IO_RQ_REPLACEMENT_FAILED</td>
<td>Number of requests failed to find space in the flash cache</td>
</tr>
<tr>
<td>FC_IO_RQ_REPLACEMENT_OLTP_FAILED</td>
<td>Number of times that client DW IOs failed to find a replacement buffer</td>
</tr>
</tbody>
</table>
Chapter 7: Monitoring Oracle Exadata System Software Components

client OLTP I/Os failed to find a replacement buffer
FC_IO_RQ_REPLACEMENT OLTP_SUCCEEDED
"Number of times that client OLTP I/Os succeeded in finding a replacement buffer"

for active secondary reads satisfied from flash cache
FC_IO_RQ_R_ACTIVE SECONDARY
FC_IO_RQ_R_ACTIVE SECONDARY_MISS
"Number of requests for active secondary reads not satisfied from flash cache"

per second for active secondary reads not satisfied from flash cache
FC_IO_RQ_R_ACTIVE SECONDARY_SEC
"Number of requests per second for active secondary reads satisfied from flash cache"

FC_IO_RQ_R_DISK_WRITER
"Number of requests read from flash cache by disk writer"
FC_IO_RQ_R_DISK_WRITER_SEC
"Number of requests per second read from flash cache by disk writer"

FC_IO_RQ_R_DW
(DW) satisfied from the flash cache
FC_IO_RQ_R_DW_SEC
(DW) per second satisfied from the flash cache
FC_IO_RQ_R_MISS
(OLTP) that did not find all data in flash cache
FC_IO_RQ_R_MISS DW
(DW) that did not find all data in flash cache
FC_IO_RQ_R_MISS DW_SEC
(DW) per second that did not find all data in flash cache
FC_IO_RQ_R_MISS_SEC
(OLTP) per second that did not find all data in flash cache
FC_IO_RQ_R_SEC
(OLTP) per second satisfied from the flash cache
FC_IO_RQ_R_SKIP
read from disk that bypass FlashCache
FC_IO_RQ_R_SKIP FC THROTTLE
FC_IO_RQ_R_SKIP FC THROTTLE_SEC
read from disk that bypass FlashCache due to heavy load on FlashCache
FC_IO_RQ_R_SKIP LG
FC_IO_RQ_R_SKIP LG_SEC
read from disk that bypass FlashCache due to the large IO size
FC_IO_RQ_R_SKIP NCMIRROR
read from disk that bypass FlashCache as the IO is on non-primary, non-active secondary mirror
FC_IO_RQ_R_SKIP SEC
FC_IO_RQ_R_W
which resulted in FlashCache being populated with data
FC_IO_RQ_R_DISK_WRITER
written to hard disks by disk writer
FC_IO_RQ_R_DISK_WRITER_SEC
per second written to hard disks by disk writer
FC_IO_RQ_R_FIRST
that are first writes into flash cache
FC_IO_RQ_R_FIRST_SEC
"Number of requests"
per second that are first writes into flash cache"

FC_IO_RQ_W_LSB
writes to flash cache due to cache hint"

FC_IO_RQ_W_LSB_DTAGE
large writes to flash cache"

FC_IO_RQ_W_LSB_MRCV
recovery large writes to flash cache"

FC_IO_RQ_W_LSB_METADATA
that are flash cache metadata writes"

FC_IO_RQ_W_LSB_METADATA_SEC
per second that are flash cache metadata writes"

FC_IO_RQ_W_LSB_OVERWRITE
that are overwrites into flash cache"

FC_IO_RQ_W_LSB_OVERWRITE_SEC
per second that are overwrites into flash cache"

FC_IO_RQ_W_LSB_POPULATE
that are population writes into flash cache due to read miss"

FC_IO_RQ_W_LSB_POPULATE_SEC
per second that are population writes into flash cache due to read miss"

FC_IO_RQ_W_LSB_SKIP
per second which resulted in FlashCache being populated with data"

FC_IO_RQ_W_SKIP_written to disk that bypass FlashCache"

FC_IO_RQ_W_SKIP_FC_THROTTLE
written to disk that bypass FlashCache due to heavy load on FlashCache"

FC_IO_RQ_W_SKIP_FC_THROTTLE_SEC
written to disk per second that bypass FlashCache due to heavy load on FlashCache"

FC_IO_RQ_W_SKIP_LG
written to disk that bypass FlashCache due to the large IO size"

FC_IO_RQ_W_SKIP_LG_SEC
written to disk per second that bypass FlashCache due to the large IO size"

FC_IO_RQ_W_SKIP_NCMIRROR
written to disk that bypass FlashCache as the IO is on non-primary, non-active secondary mirror"

FC_IO_RQ_W_SKIP_SEC
written to disk per second that bypass FlashCache"

Note the following additional details:

• Space in Exadata Smart Flash Cache is internally managed in 64K chunks known as cachelines. FC_BY_ALLOCATED represents the number of megabytes allocated in flash cache and tracks how many cachelines are used in the flash cache. If the value is close to the flash cache size, then the flash cache is fully populated.

• FC_BY_USED represents the number of megabytes used in flash cache and tracks the valid bytes in flash cache. For some workloads, like OLTP, the FC_BY_USED value can be much less than the value of FC_BY_ALLOCATED because an OLTP write might only use 8K out of a 64K cacheline. The total amount used by OLTP cachelines is tracked by FC_BY_ALLOCATED_OLTP.

• Various flash cache metrics having names ending with _DW track large read operations, which are typically associated with Data Warehouse workloads. Each of these metrics has a corresponding metric that tracks small read operations,
which are typically associated with OLTP operations. To get a total count, add the
large read metric value to the corresponding small read metric value.

For example, to get the total number of reads satisfied from flash cache, add
\texttt{FC\_IO\_RQ\_R\_DW} to \texttt{FC\_IO\_RQ\_R}.

### 7.3.1.4 What to Look For When Monitoring Exadata Smart Flash Cache

**Increased Read Latencies**

Possible issues related to Exadata Smart Flash Cache tend to be visible in the
database as increased read latencies, specifically in the \texttt{cell single block physical
read wait} event. The increased latency is usually caused when reads are issued
against the hard disks instead of being satisfied by Exadata Smart Flash Cache.

Read requests may not be satisfied from Exadata Smart Flash Cache when:

- The required data is not in Exadata Smart Flash Cache.

  In this case, requests are recorded as flash cache misses, which are visible as
  increased misses in the Flash Cache User Reads section of the AWR report. This
  is also typically accompanied by increased population writes, which are visible in
  the Flash Cache Internal Writes section of the AWR report.

  On the database, you may see a reduction in the number of \texttt{cell flash cache
  read hits compared with the physical read IO requests} or \texttt{physical read
total IO requests}.

- The data is not eligible for Exadata Smart Flash Cache.

  To maximize cache efficiency, when an I/O request is sent from the database to
  the storage servers, it includes a hint that indicates whether or not the data should
  be cached in Exadata Smart Flash Cache.

  If data is not eligible for caching, the corresponding I/O is visible in the Flash
  Cache User Reads - Skips section of the AWR report, along with the reason why
  the read was not eligible. Possible reasons include:

  - The grid disk caching policy is set to \texttt{none}.
  - The database segment is configured with the \texttt{CELL\_FLASH\_CACHE NONE} storage
    option.
  - The I/O type and situation precludes caching. For example, the I/O is related
to an RMAN backup operation and the hard disks are not busy. If this is a
common occurrence, it will be evident in Top IO Reasons section of the AWR
report

  In some cases, the \texttt{cell single block physical read} latency may increase, with no
  apparent difference in Exadata Smart Flash Cache behavior. This can be caused by
  increased I/O load, especially on hard disks.

  Occasional long latencies in \texttt{cell single block physical read}, usually visible as
  a long tail in the \texttt{cell single block physical read} histogram may simply indicate
  that the data is read from hard disk, and does not necessarily indicate a problem with
  Exadata Smart Flash Cache.

**Skipped Writes**

When using Exadata Smart Flash Cache in Write-Back mode, most database write
requests are absorbed by the cache. However, some write requests may skip the
cache. The most common reason for skipping the cache is when the request writes a large amount of data that will be read once, such as temp sorts, or it is not expected to be read at all in the foreseeable future, such as backups and archiving.

Another common reason for skipping large writes is Disk Not Busy, which typically means that there is no benefit to using Exadata Smart Flash Cache because the hard disks have sufficient capacity to handle the write requests.

If skipping Exadata Smart Flash Cache for large writes is causing a performance issue, then it is typically visible in the database with corresponding long latencies for the direct path write or direct path write temp wait events.

The reasons for rejecting large writes are visible in the Flash Cache User Writes - Large Write Rejections section of the AWR report.

Database Working Set Size

The database working set refers to the subset of most commonly accessed information in the database. In most cases, the database working set is fairly stable. However, if for any reason, the working set does not fit in Exadata Smart Flash Cache, you may see the following symptoms:

- Increased cache misses, indicating that data is not in Exadata Smart Flash Cache. This is visible in the Flash Cache User Reads section of the AWR report, or the \texttt{FC\textunderscore IO\textunderscore R\textunderscore R\textunderscore MISS\textunderscore SEC} cell metric.
- Increased population activity to populate data not in Exadata Smart Flash Cache. This is visible in the Flash Cache Internal Writes section of the AWR report, or the \texttt{FC\textunderscore IO\textunderscore [BY\textbar RQ]\textunderscore W\textunderscore POPULATE\textunderscore SEC} cell metrics.
- Increased disk writer activity, indicating that dirty cachelines have to be written to disk, so that the cacheline can be reused to cache other data. Disk writer activity is visible in the Flash Cache Internal Reads section of the AWR report, or the \texttt{FC\textunderscore IO\textunderscore [RQ\textbar BY]\textunderscore [R\textbar W]\textunderscore DISK\textunderscore WRITER\textunderscore SEC} cell metrics.
- Increased first writes, indicating new data is being written to Exadata Smart Flash Cache. A large number of first writes with few overwrites means new data is being written to Exadata Smart Flash Cache. This is visible in the Flash Cache User Writes section of the AWR report, or in the \texttt{FC\textunderscore IO\textunderscore [RQ\textbar BY]\textunderscore W\textunderscore FIRST\textunderscore SEC} cell metrics.

In this case:

- Review the database access patterns for tuning opportunities that reduce the amount of data being accessed.
- Consider increasing the number of available storage servers to deliver more space for Exadata Smart Flash Cache.
- Review your I/O Resource Management (IORM) quotas for Exadata Smart Flash Cache and allocate space where it is most required.
- Consider using Extreme Flash storage servers to eliminate the disk I/Os.

Miscellaneous Issues

There are some cases when the increased cell single block physical read latency may not be due to cell performance, but may be caused by something else along the IO path, such as network contention or contention for database server CPU resources.
The histograms for cell single block physical read and small reads are available in the Single Block Reads and Small Read Histogram - Detail sections of the AWR report under Exadata Statistics > Performance Summary. The cell single block physical read histogram shows latencies measured by Oracle Database, while the small read histograms show latencies measured in the storage server.

A histogram with a significant number of occasional long latencies is said to have a long tail. When the histograms for cell single block physical read and small reads have long tails, then this is an indication of slow read times in the storage server, which would warrant further investigation of the other I/O performance statistics. See Monitoring Cell Disk I/O.

If the cell single block physical read histogram has a long tail that is not present in the small read histograms, then the cause is generally not in the storage server, but rather something else in the I/O path, such as bottlenecks in the network or contention for compute node CPU.

**Related Topics**
- Faster Performance for Large Analytic Queries and Large Loads

### 7.3.2 Monitoring PMEM Cache

Persistent Memory (PMEM) cache provides direct access to persistent memory on the storage servers using Remote Direct Memory Access (RDMA), enabling lower read latencies and faster response times. PMEM is a persistent memory tier that was first released on Exadata X8M. When database clients read from the PMEM cache, the client software performs an RDMA read of the cached data, which bypasses the storage server software and results in much lower read latencies.

PMEM cache works in conjunction with Exadata Smart Flash Cache. If available, data that is not in PMEM cache may be retrieved from Exadata Smart Flash Cache. Similarly, as data is written out of PMEM cache, it is written to Exadata Smart Flash Cache when using Write-Back mode.

Statistics from PMEM cache are slightly different than the rest of the Exadata statistics. Because clients issue RDMA I/O directly to PMEM cache, the request does not go to cellsrv, so the storage server cannot account for the RDMA I/Os. For this reason, there are no cell metrics for PMEM cache I/O. Instead, Oracle Database statistics account for the I/O that is performed using RDMA.

Performance issues related to PMEM cache typically cause latency increases in the Oracle Database cell single block physical read wait event. However, bear in mind that Exadata Smart Flash Cache is still available to service the requests, and although requests from Exadata Smart Flash Cache generally experience higher read latencies compared to PMEM cache, the requests still benefit from the fast I/O provided by flash.

- Monitoring PMEM Cache Using AWR
- Monitoring PMEM Cache Using Database Statistics and Wait Events
- Monitoring PMEM Cache Using Exadata Metrics
- What to Look For When Monitoring PMEM Cache
7.3.2.1 Monitoring PMEM Cache Using AWR

Automatic Workload Repository (AWR) contains information relating to PMEM cache. Following are descriptions and examples of the most commonly used sections in the AWR report that contain information about PMEM cache. By reviewing these sections of the AWR report, administrators can understand the operation of PMEM cache.

Note that because storage servers cannot account for the RDMA I/Os, the Oracle Database statistics are critical in understanding the use of PMEM cache. The PMEM cache statistics in the Exadata-specific sections of the AWR report only include I/Os that as serviced by cellsrv, which are not RDMA I/Os.

Apart from the sections described below, the AWR report also contains sections for PMEM Cache User Writes, and PMEM Cache Internal Reads. These sections are not described in detail because they relate to the use of PMEM cache in Write-Back mode, which is not generally recommended.

**Database IOs**

A summary of Database IOs may be found in the Single Block Reads section of the AWR report. This summary contains information about the effectiveness of PMEM cache. The Single Block Reads section is located in the AWR report under Exadata Statistics > Performance Summary.

In the following example, the overwhelming majority of all read I/O requests (approximately 88%) are serviced using RDMA reads to PMEM cache (cell RDMA reads), at a rate of more than 176,000 per second. Nearly all of the rest are satisfied using non-RDMA PMEM cache reads (cell pmem cache read hits), while the remaining reads are serviced by using Exadata Smart Flash Cache (cell flash cache read hits). Notice the massive difference in throughput for each different I/O type, which illustrates the power of PMEM cache and RDMA.

<table>
<thead>
<tr>
<th>Database IOs</th>
<th>Value</th>
<th>per Sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>physical read total IO requests</td>
<td>393,579,196</td>
<td>200,601.02</td>
</tr>
<tr>
<td>physical read IO requests</td>
<td>393,571,051</td>
<td>200,596.87</td>
</tr>
<tr>
<td>cell flash cache read hits</td>
<td>183,893</td>
<td>93.73</td>
</tr>
<tr>
<td>cell ram cache read hits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cell pmem cache read hits</td>
<td>46,551,898</td>
<td>23,726.76</td>
</tr>
<tr>
<td>cell RDMA reads</td>
<td>346,841,370</td>
<td>176,779.50</td>
</tr>
</tbody>
</table>

**PMEM Cache Configuration and Space Usage**

The PMEM Cache Configuration section contains summary information including the caching mode (Write-Through or Write-Back) and overall size. The PMEM Cache Space Usage section provides summary statistics on space usage in PMEM cache.

The following example shows PMEM cache configured in Write-Through mode on all cells, with a total size of approximately 1509 GB.


**Figure 7-11   AWR Report: PMEM Cache Configuration**

**PMEM Cache Configuration**

- These statistics are collected by the cells and are not restricted to this database or instance

<table>
<thead>
<tr>
<th>Mode</th>
<th>Size (GB)</th>
<th>Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>write-through</td>
<td>1,509.56</td>
<td>All</td>
</tr>
</tbody>
</table>

The following example of the PMEM Cache Space Usage section shows 172 GB of PMEM cache spread evenly across 3 cells.

**Figure 7-12   AWR Report: PMEM Cache Space Usage**

**PMEM Cache Space Usage**

- These statistics are collected by the cells and are not restricted to this database or instance
- Space is at the time of the end snapshot
- Ordered by Space (GB) desc

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>Space Allocated</th>
<th>% OLTP</th>
<th>%Clean</th>
<th>%Dirty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (3)</td>
<td>172.00 GB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm0celadm11</td>
<td>57.35 GB</td>
<td>87.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm0celadm12</td>
<td>57.35 GB</td>
<td>87.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm0celadm10</td>
<td>57.30 GB</td>
<td>87.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PMEM Cache User Reads**

The PMEM Cache User Reads section shows information about read requests, read throughput, and read efficiency. Because storage servers cannot account for RDMA I/Os, the statistics only relate to non-RDMA reads processed by `cellsrv`. Hits are for I/Os satisfied using PMEM cache, while misses indicate that the data was not in PMEM cache.

**Figure 7-13   AWR Report: PMEM Cache User Reads**

**PMEM Cache User Reads**

- These statistics are collected by the cells and are not restricted to this database or instance
- ordered by hits desc

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>Hits</th>
<th>Hits per Sec</th>
<th>%Hit</th>
<th>Misses</th>
<th>Misses per Sec</th>
<th>Hits MB</th>
<th>Hits MB/s</th>
<th>Misses MB</th>
<th>Misses MB/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (3)</td>
<td>46,495,226</td>
<td>23,697.87</td>
<td>99.60</td>
<td>185,811</td>
<td>94.70</td>
<td>182,188.40</td>
<td>92.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm0celadm10</td>
<td>15,578,238</td>
<td>7,939.98</td>
<td>99.66</td>
<td>53,455</td>
<td>27.25</td>
<td>61,056.54</td>
<td>31.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm0celadm12</td>
<td>15,556,409</td>
<td>7,928.85</td>
<td>99.66</td>
<td>52,506</td>
<td>26.76</td>
<td>60,936.96</td>
<td>31.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dbm0celadm11</td>
<td>15,360,581</td>
<td>7,829.04</td>
<td>99.48</td>
<td>79,850</td>
<td>40.70</td>
<td>60,194.90</td>
<td>30.68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PMEM Cache Internal Writes
The PMEM Cache Internal Writes section shows writes to PMEM cache that are performed by Oracle Exadata System Software. The internal writes are I/Os that populate PMEM cache.

**Figure 7-14   AWR Report: PMEM Cache Internal Writes**

### 7.3.2.2 Monitoring PMEM Cache Using Database Statistics and Wait Events

The following table describes database statistics that are useful for monitoring PMEM cache. The statistics are available in various dynamic performance views, including V$SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell RDMA reads</td>
<td>The number of PMEM cache read requests using RDMA</td>
</tr>
<tr>
<td>cell RDMA writes</td>
<td>The number of PMEM cache write requests using RDMA</td>
</tr>
<tr>
<td>cell pmem cache read hits</td>
<td>The number of non-RDMA read requests processed by <code>cellsrv</code> resulting in a PMEM cache hit</td>
</tr>
<tr>
<td>cell pmem cache writes</td>
<td>The number of non-RDMA write requests processed by <code>cellsrv</code> resulting in a PMEM cache write</td>
</tr>
</tbody>
</table>

The database wait events that are useful for monitoring PMEM cache are the same wait events that are associated with Exadata Smart Flash Cache.

The use of PMEM cache is typically indicated by extremely low latency for the cell single block physical read event, especially for operations using RDMA.

The following table describes database wait events that are useful for monitoring PMEM cache. The wait events are visible in various dynamic performance views, including V$SESSION, V$SYSTEM_EVENT and V$SESSION_EVENT, and may be displayed in the Wait Event sections of the AWR report.
### Wait Event

<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell list of blocks physical read</td>
<td>This wait event occurs during recovery or during buffer pre-fetching (rather than performing multiple single-block reads). It is used to monitor database blocks that must be changed as part of recovery and are read in parallel for the database. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the number of blocks processed during the I/O read operation. This wait event is equivalent to db file parallel read for a cell.</td>
</tr>
<tr>
<td>cell multiblock physical read</td>
<td>The wait time is the actual time it takes to do all of the I/Os when reading multiple data blocks. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the total number of bytes processed during the I/O read operation. This wait event is equivalent to db file scattered read for a cell.</td>
</tr>
<tr>
<td>cell single block physical read</td>
<td>The wait time is the actual time it takes to do the I/O while a sequential read for the database is performed. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the number of bytes processed during the I/O read operation. This wait event is equivalent to db file sequential read for a cell.</td>
</tr>
<tr>
<td>cell interconnect retransmit during physical read</td>
<td>This wait event appears during retransmission for an I/O of a single-block or multiblock read. The cell hash number in the P1 column in the V$SESSION_WAIT view is the same cell identified for cell single block physical read and cell multiblock physical read. The P2 column contains the subnet number to the cell, and the P3 column contains the number of bytes processed during the I/O read operation.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.

#### 7.3.2.3 Monitoring PMEM Cache Using Exadata Metrics

Exadata metrics that are related to PMEM cache are identified in the Exadata storage server METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=PMEMCACHE.
Note that because reads from PMEM cache are primarily performed using RDMA calls from Oracle Database, there are no Exadata metrics that tally PMEM cache I/O.

Example 7-2  Displaying PMEM Cache Metric Definitions

This example shows how to display the PMEM cache metric definitions that are available in the Oracle Exadata System Software.

```
CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE
OBJECTTYPE = PMEMCACHE

PC_BY_ALLOCATED         "Number of megabytes allocated in PMEM cache"
```

Note the following additional details:

- **PC_BY_ALLOCATED** represents the number of megabytes allocated in PMEM cache and tracks how many cachelines are used in the PMEM cache. If the value is close to the PMEM cache size, then the PMEM cache is fully populated.

  This metric is also available per database (**DB_PC_BY_ALLOCATED**), and per PDB (**PDB_PC_BY_ALLOCATED**).

7.3.2.4 What to Look For When Monitoring PMEM Cache

**Database Working Set**

Performance is maximized if the entire database working set fits into PMEM cache. However, the PMEM cache is much smaller than Exadata Smart Flash Cache and it is likely that the database working set, or even the most frequently accessed portion of it, might not reside in PMEM cache.

If the size of the working set increases you may observe more reads being satisfied by Exadata Smart Flash Cache or hard disk instead of PMEM cache. This is magnified when access patterns on the data are more random. Consequently, you may observe increased read latencies, specifically in the **cell single block physical read wait** event.

**Non-RDMA Reads**

PMEM cache is most effective when it is used in conjunction with RDMA. However, in various situations requests to **cellsrv** might still be satisfied using PMEM cache. In such cases, you will see the associated reads as **cell pmem cache read hits** rather than **cell RDMA reads**.

A non-RDMA read may occur for the following reasons:

- Based on the type of data being read, Oracle Database may not be able to perform an RDMA read. Direct reads and reads from the Oracle Database control file are examples of this. To identify the cause you may be need to correlate with other statistics or wait events. For example, significant activity against the control files is usually accompanied by the **control file sequential read** wait event, or visible in the IOStat by Filetype section of the AWR report.

- In the Oracle Database client, buffers may not be registered for RDMA. This normally occurs shortly after a client process starts or while PMEM cache is populating.
- Inside PMEM cache, the RDMA hash table may not contain the required metadata (or it may be marked as invalid). This normally occurs while PMEM cache is populating.
- An RDMA read times out after exceeding the default lease time (20 ms).
- Memory limitations prevent the creation of memory structures that are required for RDMA.
- An RDMA read error occurs.

7.3.3 Monitoring Exadata Smart Flash Log

Exadata Smart Flash Log reduces the average latency for performance-sensitive redo log write I/O operations, thereby eliminating performance bottlenecks that may occur due to slow redo log writes. Exadata Smart Flash Log removes latency spikes by simultaneously performing redo log writes to two media devices. The redo write is acknowledged as soon as the first write to either media device completes.

Originally, Exadata Smart Flash Log was used to perform simultaneous writes to disk and flash storage. However, Oracle Exadata System Software release 20.1 adds a further optimization, known as Smart Flash Log Write-Back, that uses Exadata Smart Flash Cache in Write-Back mode instead of disk storage.

- Monitoring Exadata Smart Flash Log Using AWR
- Monitoring Exadata Smart Flash Log Using Database Statistics and Wait Events
- Monitoring Exadata Smart Flash Log Using Exadata Metrics
- What to Look For When Monitoring Exadata Smart Flash Log

7.3.3.1 Monitoring Exadata Smart Flash Log Using AWR

Following are descriptions and examples of the Flash Log Statistics sections in the AWR report. By reviewing these sections of the AWR report, administrators can understand the operation of Exadata Smart Flash Log.

Flash Log

The Flash Log section contains summary statistics about the operation of Exadata Smart Flash Log, including the number and of writes to Exadata Smart Flash Log, the number outliers prevented by Exadata Smart Flash Log, and the number of times when Exadata Smart Flash Log is skipped.

The following show an example of the Flash Log section. As shown in the example, when Exadata Smart Flash Log operates optimally, I/O is evenly distributed across all of the cells and there are no skips.
If the Skip Count is greater than zero, then the Flash Log Skip Details section contains the reasons for skipping Exadata Smart Flash Log.

Redo Write Histogram

The Redo Write Histogram section provides histograms that show the log file parallel write latency from the database, along with the redo write request completion latency from the storage servers. By comparing the database and cell latency histograms, you can determine if the high latency outliers are related to processing bottlenecks on the storage servers.
7.3.3.2 Monitoring Exadata Smart Flash Log Using Database Statistics and Wait Events

The following table describes various database statistics that are useful for monitoring redo write and Exadata Smart Flash Log performance. The statistics are available in various dynamic performance views, including V$SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>redo write size count (4KB)</td>
<td>Number of redo writes where the redo write size is smaller than the size indicated in parenthesis. For example, redo write size count (4KB) includes redo writes smaller than 4 KB in size, and redo write size count (inf) includes redo writes over 1 MB in size.</td>
</tr>
<tr>
<td>redo write size count (8KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (16KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (32KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (64KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (128KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (256KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (512KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (1024KB)</td>
<td></td>
</tr>
<tr>
<td>redo write size count (inf)</td>
<td></td>
</tr>
<tr>
<td>redo writes</td>
<td>Total number of writes to the redo log files by the Oracle Database log writer (LGWR) process(es).</td>
</tr>
</tbody>
</table>

The following table describes database wait events that are useful for monitoring redo write and Exadata Smart Flash Log performance. The wait events are visible in various dynamic performance views, including V$SESSION, V$SYSTEM_EVENT and V$SESSION_EVENT, and may be displayed in the Wait Event sections of the AWR report.

<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>log file parallel write</td>
<td>The Oracle Database log writer (LGWR) process waits on this event when it is waiting for the completion of writes to the redo log file. Efficient use of Exadata Smart Flash Log is indicated by comparatively lower time waiting on this event.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.

7.3.3.3 Monitoring Exadata Smart Flash Log Using Exadata Metrics

Exadata metrics that are related to Exadata Smart Flash Log provide information about flash log utilization, such as the number of megabytes written per second. Flash Log metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=FLASHLOG.
Example 7-3  Displaying Flash Log Metric Definitions

This example shows how to display the flash log metric definitions that are available in
the Oracle Exadata System Software.

CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE
OBJECTTYPE = FLASHLOG

FL_ACTUAL_OUTLIERS  "The number of times redo
writes to flash and disk both exceeded the outlier threshold"

FL_BY_KEEP  "The amount of redo data
saved on flash due to disk I/O errors"

FL_DISK_FIRST  "The number of times redo
writes first completed to disk"

FL_DISK_IO_ERRS  "The number of disk I/O
errors encountered by Smart Flash Logging"

FL_EFFICIENCY_PERCENTAGE  "The efficiency of Smart Flash
Logging expressed as a percentage"

FL_EFFICIENCY_PERCENTAGE_HOUR  "The efficiency of Smart Flash
Logging over the last hour expressed as a percentage"

FL_FLASH_FIRST  "The number of times redo
writes first completed to flash"

FL_FLASH_IO_ERRS  "The number of flash I/O
errors encountered by Smart Flash Logging"

FL_FLASH_ONLY_OUTLIERS  "The number of times redo
writes to flash exceeded the outlier threshold"

FL_IO_DB_BY_W  "The number of MB written to
hard disk by Smart Flash Logging"

FL_IO_DB_BY_W_SEC  "The rate which is the number
of MB per second written to hard disk by
Smart Flash Logging"

FL_IO_FL_BY_W  "The number of MB written to
flash by Smart Flash Logging"

FL_IO_FL_BY_W_SEC  "The rate which is the number
of MB per second written to flash by Smart Flash Logging"

FL_IO_TM_W  "Cumulative latency of all
redo log writes"

FL_IO_TM_W_RQ  "Average latency of all redo
log writes"

FL_IO_W  "The number of writes serviced
by Smart Flash Logging"

FL_IO_W_SKIP_BUSY  "The number of redo writes
that could not be serviced by Smart Flash Logging because too much data
had not yet been written to disk"

FL_IO_W_SKIP_BUSY_MIN  "The number of redo writes
during the last minute that could not be serviced by Smart Flash
Logging because too much data had not yet been written to disk"

FL_IO_W_SKIP_DISABLED_GD  "The number of redo writes
that could not be serviced by Smart Flash Logging because it was
disabled for the redo log's grid disk"

FL_IO_W_SKIP_IORM_LIMIT  "The number of redo writes
that could not be serviced by Smart Flash Logging because the IORM
limit had been reached for the redo log's grid disk"

FL_IO_W_SKIP_IORM_PLAN  "The number of redo writes
that could not be serviced by Smart Flash Logging because it was
disabled by the IORM plan"

FL_IO_W_SKIP_LARGE  "The number of large redo
writes..."
writes that could not be serviced by Smart Flash Logging because the size of the data was larger than the amount of available space on any flash disk"

\[
\text{FL\_IO\_W\_SKIP\_LOG\_ON\_FLASH} \quad \text{"The number of redo writes that could not be serviced by Smart Flash Logging because the redo log resided on flash"}
\]

\[
\text{FL\_IO\_W\_SKIP\_NO\_BUFFER} \quad \text{"The number of redo writes that could not be serviced by Smart Flash Logging because of lack of available buffers"}
\]

\[
\text{FL\_IO\_W\_SKIP\_NO\_FL\_DISKS} \quad \text{"The number of redo writes that could not be serviced by Smart Flash Logging because there were no available Flash Log disks"}
\]

\[
\text{FL\_PREVENTED\_OUTLIERs} \quad \text{"The number of times redo writes to disk exceeded the outlier threshold; these would have been outliers had it not been for Smart Flash Logging"}
\]

\[
\text{FL\_RQ\_TM\_W} \quad \text{"Cumulative latency of all redo log write requests (includes network and other processing overhead)"}
\]

\[
\text{FL\_RQ\_TM\_W\_RQ} \quad \text{"Average latency of all redo log write requests"}
\]

\[
\text{FL\_RQ\_W} \quad \text{"The number of redo log write requests serviced (includes requests which were not handled by Smart Flash Logging)"
\]

\[
\text{FL\_SKIP\_OUTLIERs} \quad \text{"The number of times redo writes to disk exceeded the outlier threshold when Smart Flash Logging was not used"}
\]

Note the following additional details:

- **FL\_RQ\_TM\_W** tracks the cumulative redo log write request latency, which includes networking and other overhead. To determine the overhead component, get the latency overhead due to factors such as network and processing, you can use subtract FL\_IO\_TM\_W from FL\_RQ\_TM\_W.

- **FL\_SKIP\_OUTLIERs** tracks the number of outliers when redo log write skips use of the flash log. An outlier is a redo log write that exceeds 0.5 seconds.

The use of the flash log can be skipped because of the reasons associated with FL\_IO\_W\_SKIP\_DISABLED\_GD, FL\_IO\_W\_SKIP\_IORM\_LIMIT, FL\_IO\_W\_SKIP\_IORM\_PLAN, FL\_IO\_W\_SKIP\_LOG\_ON\_FLASH, or FL\_IO\_W\_SKIP\_NO\_FL\_DISKS.

- For simultaneous writes to disk and flash storage, the disk controller write cache can absorb some writes quicker than flash. Consequently, it is normal for a significant proportion of redo log write operations to complete to disk before flash. In some cases, the value of FL\_DISK\_FIRST can exceed the value of FL\_FLASH\_FIRST. However, this does not mean that the Exadata Smart Flash Logging feature is ineffective or unnecessary.

### 7.3.3.4 What to Look For When Monitoring Exadata Smart Flash Log

#### General Performance

Performance issues related to redo logging typically exhibit high latency for the log file sync wait event in the Oracle Database user foreground processes, with corresponding high latency for log file parallel write in the Oracle Database log.
writer (LGWR) process. Because of the performance-critical nature of redo log writes, occasional long latencies for log file parallel write may cause fluctuations in database performance, even if the average log file parallel write wait time is acceptable.

If any of these are occurring, then it may be indicative of an issue with Exadata Smart Flash Log performance.

**Redo Write Histograms**

The log file parallel write wait event indicates the amount of time that the database waits on a redo log write. The log file parallel write histogram shows the number of times the redo write completed within a specified time range. Similarly, the redo log write completions statistic indicates the amount of time that the storage server spends processing redo write requests, and the redo log write completions histogram shows the number of times the redo write request was completed within a specified time range. Both histograms are shown in the Redo Write Histogram section of the AWR report.

A histogram with a significant number of occasional long latencies is said to have a long tail. When both of the histograms in the Redo Write Histogram section of the AWR report have long tails, then this is an indication of slow write times on the storage server, which would warrant further investigation of the other I/O performance statistics. See Monitoring Cell Disk I/O.

If the log file parallel write histogram has a long tail that is not present in the redo log write completions histogram, then the cause is generally not in the storage server, but rather something else in the I/O path, such as bottlenecks in the network or contention for compute node CPU.

**Skipping**

Increased redo write latencies can also occur when the Exadata Smart Flash Log is skipped, and the redo write goes only to disk. Both the AWR report and the storage server metrics show the number of redo log writes that skipped Exadata Smart Flash Log. Skipping may occur when Exadata Smart Flash Log contains too much data that has not yet been written to disk.

There are a few factors that can cause redo writes to skip Exadata Smart Flash Log:

- **Flash disks with high write latencies.**
  This can be observed in various IO Latency tables located in the Exadata Resource Statistics section of the AWR report, and in the Exadata CELLDISK metrics. This can also be identified by checking the FL_FLASH_ONLY_OUTLIERS metric. If the metric value is high, this indicates a flash disk performance issue.

- **Hard disks with high latencies or high utilization.**
  Prior to Oracle Exadata System Software release 20.1, redo log writes are written to both Exadata Smart Flash Log and hard disk. If the hard disks experience high latencies or high utilization, redo log write performance can be impacted.

  This can be observed in various IO Latency tables located in the Exadata Resource Statistics section of the AWR report, and in the Exadata CELLDISK metrics. This can also be identified by checking the Outliers columns in the Flash Log section of the AWR report, or the FL_PREVENTED_OUTLIERS storage server metric. A large number of prevented outliers may indicate that the hard disk writes are taking a long time.
In this case, although Exadata Smart Flash Log prevents outliers, overall throughput may be limited due to the queue of redo log data that must be written to disk.

Oracle Exadata System Software release 20.1 adds a further optimization, known as Smart Flash Log Write-Back, that uses Exadata Smart Flash Cache in Write-Back mode instead of disk storage, thereby eliminating the hard disks as a potential performance bottleneck. Depending on the system workload, this feature can improve overall log write throughput by up to 250%.

7.3.4 Monitoring PMEM Log

Redo log writes are critical database operations and need to complete in a timely manner to prevent load spikes or stalls. Exadata Smart Flash Log is designed to prevent redo write latency outliers. PMEM Log helps to further reduce redo log write latency by using Persistent Memory (PMEM) and Remote Direct Memory Access (RDMA).

With PMEM Log, database clients send I/O buffers directly to PMEM on the storage servers using RDMA, thereby reducing transport latency. The cell server (cellsrv) then writes the redo to Exadata Smart Flash Log (if enabled) and disk at a later time.

Reduced redo log write latency improves OLTP performance, resulting in higher transaction throughput. In cases where PMEM Log is bypassed, Exadata Smart Flash Log can still be used.

- Monitoring PMEM Log Using Database Statistics and Wait Events
- What to Look For When Monitoring PMEM Log

7.3.4.1 Monitoring PMEM Log Using Database Statistics and Wait Events

The following table describes database statistics that are useful for monitoring PMEM Log performance. The statistics are available in various dynamic performance views, including `V$SYSSTAT`, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell pmem log writes</td>
<td>The number of redo log write requests that used PMEM log</td>
</tr>
</tbody>
</table>

The following table describes database wait events that are useful for monitoring PMEM Log performance. The wait events are visible in various dynamic performance views, including `V$SESSION`, `V$SYSTEM_EVENT` and `V$SESSION_EVENT`, and may be displayed in the Wait Event sections of the AWR report.

<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>log file parallel write</td>
<td>The Oracle Database log writer (LGWR) process waits on this event when it is waiting for the completion of writes to the redo log file. Efficient use of PMEM Log is indicated by comparatively lower time waiting on this event.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.
7.3.4.2 What to Look For When Monitoring PMEM Log

**General Performance**

Performance issues related to redo logging typically exhibit high latency for the log file sync wait event in the Oracle Database user and foreground processes, with corresponding high latency for log file parallel write in the Oracle Database log writer (LGWR) process. Because of the performance-critical nature of redo log writes, occasional long latencies for log file parallel write may cause fluctuations in database performance, even if the average log file parallel write wait time is acceptable.

If any of these are occurring, then it may be indicative of an issue with PMEM Log performance.

**Bypassing PMEM Log**

Increased redo write latencies when using PMEM Log can occur when PMEM Log is bypassed. When PMEM Log is bypassed, the request is sent to the cellsrv, and Exadata Smart Flash Log is still used (if available). However, when the bypass request is sent, it has to ensure that there is no conflict with a previous PMEM Log request. This conflict checking, which requires scanning PMEM Log, makes bypass writes more expensive to process, and can result in higher than expected redo log write latencies.

There are several possible causes that result in a small number of PMEM Log bypasses. Under normal circumstances, the number of bypasses should be substantially less than 1% of the total number of PMEM Log requests. A high number of PMEM Log bypasses is likely to be a symptom of another problem, such as congestion on the RoCE Network Fabric.

7.3.5 Monitoring Smart I/O

Exadata Smart Scan offloads data search and retrieval processing to the Exadata storage servers. It is able to evaluate database predicates on the storage server to optimize efficiency and performance, especially for large queries and certain classes of bulk data processing.

Smart Scan also uses storage indexes, which are maintained automatically on the storage servers, to further optimize filtering by eliminating unnecessary I/O.

Columnar cache is another smart I/O optimization. The columnar cache is a section of Exadata Smart Flash Cache that stores data in columnar format. The columnar cache is automatically populated and maintained by the storage server when a Smart Scan is performed.

- Monitoring Smart I/O Using AWR
- Monitoring Smart I/O Using Database Statistics and Wait Events
- Monitoring Smart I/O Using SQL Monitor
- Monitoring Smart I/O Using SQL Explain Plan
- Monitoring Smart I/O Using Exadata Metrics
- What to Look For When Monitoring Smart I/O
Related Topics

- **Offloading of Data Search and Retrieval Processing**
  One of the most powerful features of Oracle Exadata System Software is that it offloads the data search and retrieval processing to the storage servers.

- **Storage Index**
  Oracle Exadata Storage Servers maintain a storage index which contains a summary of the data distribution on the disk.

- **In-Memory Columnar Format Support**
  In an Oracle Exadata Database Machine environment, the data is automatically stored in In-Memory columnar format in the flash cache when it will improve performance.

### 7.3.5.1 Monitoring Smart I/O Using AWR

The Smart IO section of the AWR report contains summarized statistical information for Smart I/O operations on the whole system and on each storage server. In particular, it tracks the amount of I/O that is eligible for optimization using Smart Scan, the savings from using storage indexes, the amount of data in Exadata Smart Flash Cache, and the amount of data filtered by the storage servers without transportation back to the database server. It also shows information about passthru bytes and reverse offload.

#### Figure 7-17  AWR Report: Smart IO

<table>
<thead>
<tr>
<th>MB Requested</th>
<th>Storage Index</th>
<th>Flash Cache</th>
<th>Offload</th>
<th>Passthru</th>
<th>Reverse Offload</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MB</td>
<td>% Optimized</td>
<td>MB</td>
<td>% Optimized</td>
</tr>
<tr>
<td>Total (14)</td>
<td>1,880,672.25</td>
<td>45,870.06</td>
<td>1,177,631.25</td>
<td>62.62</td>
<td>119.38</td>
</tr>
<tr>
<td>db00ceidm10</td>
<td>7.17</td>
<td>134,838.75</td>
<td>3,288.75</td>
<td>84,433.00</td>
<td>62.62</td>
</tr>
<tr>
<td>db00ceidm03</td>
<td>7.18</td>
<td>134,745.44</td>
<td>3,288.47</td>
<td>84,326.44</td>
<td>62.58</td>
</tr>
<tr>
<td>db00ceidm02</td>
<td>7.16</td>
<td>134,629.53</td>
<td>3,288.43</td>
<td>83,969.09</td>
<td>62.08</td>
</tr>
<tr>
<td>db00ceidm01</td>
<td>7.15</td>
<td>134,559.80</td>
<td>3,281.90</td>
<td>84,546.61</td>
<td>62.83</td>
</tr>
<tr>
<td>db00ceidm08</td>
<td>7.15</td>
<td>134,496.41</td>
<td>3,280.40</td>
<td>83,952.23</td>
<td>62.42</td>
</tr>
<tr>
<td>db00ceidm12</td>
<td>7.14</td>
<td>134,363.95</td>
<td>3,277.17</td>
<td>84,433.67</td>
<td>62.84</td>
</tr>
<tr>
<td>db00ceidm11</td>
<td>7.14</td>
<td>134,333.86</td>
<td>3,276.43</td>
<td>83,961.96</td>
<td>62.50</td>
</tr>
<tr>
<td>db00ceidm13</td>
<td>7.14</td>
<td>134,235.08</td>
<td>3,274.03</td>
<td>83,623.67</td>
<td>62.32</td>
</tr>
<tr>
<td>db00ceidm09</td>
<td>7.13</td>
<td>134,172.17</td>
<td>3,272.49</td>
<td>84,051.25</td>
<td>62.64</td>
</tr>
<tr>
<td>db00ceidm04</td>
<td>7.13</td>
<td>134,136.84</td>
<td>3,271.63</td>
<td>84,305.92</td>
<td>62.85</td>
</tr>
<tr>
<td>db00ceidm07</td>
<td>7.13</td>
<td>134,116.95</td>
<td>3,271.15</td>
<td>83,795.31</td>
<td>62.48</td>
</tr>
<tr>
<td>db00ceidm06</td>
<td>7.13</td>
<td>134,106.52</td>
<td>3,270.88</td>
<td>84,152.72</td>
<td>62.72</td>
</tr>
<tr>
<td>db00ceidm14</td>
<td>7.13</td>
<td>134,103.13</td>
<td>3,270.81</td>
<td>83,967.59</td>
<td>62.61</td>
</tr>
<tr>
<td>db00ceidm05</td>
<td>7.12</td>
<td>133,843.03</td>
<td>3,264.46</td>
<td>84,511.08</td>
<td>63.14</td>
</tr>
</tbody>
</table>

Statistics relating to columnar cache are contained in the sections of the AWR report that contain information about Exadata Smart Flash Cache. In particular, note the following sections in the AWR report:

- **Flash Cache Space Usage** — shows the amount of space used for columnar cache
- **Flash Cache User Reads** — shows statistics for reads from columnar cache
• Flash Cache Internal Writes — shows statistics for population writes into columnar cache
• Columnar Cache Efficiency — shows statistics for reads from columnar cache, for each columnar cache format
• Columnar Cache Population — shows statistics for population writes into columnar cache, for each columnar cache format
• Columnar Cache Population (Current) — shows the size of columnar cache areas, for each columnar cache format, including outstanding columnar cache population jobs

Related Topics
• Monitoring Exadata Smart Flash Cache Using AWR

7.3.5.2 Monitoring Smart I/O Using Database Statistics and Wait Events

The following table describes various database statistics that are useful for monitoring smart I/O operations. The statistics are available in various dynamic performance views, including V$SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell IO uncompressed bytes</td>
<td>The total size of uncompressed data that is processed on the cell. For operations on segments compressed using Exadata Hybrid Columnar Compression, this statistic is the size of the data after decompression.</td>
</tr>
<tr>
<td>cell num bytes in passthru due to quarantine</td>
<td>The number of bytes that were not offloaded and sent back to the database for processing due to a quarantine on the cell.</td>
</tr>
<tr>
<td>cell num bytes in passthru during predicate offload</td>
<td>The number of bytes that were not offloaded and sent back to the database for processing.</td>
</tr>
<tr>
<td>cell physical IO bytes added to storage index</td>
<td>The number of bytes added to the storage index during a Smart Scan. This is an indication that the storage index is being built.</td>
</tr>
<tr>
<td>cell physical IO bytes eligible for predicate offload</td>
<td>The number of bytes on-disk eligible for predicate offload. For example, when using columnar cache, this is the size of columnar cache instead of the on-disk size.</td>
</tr>
<tr>
<td>cell physical IO bytes eligible for smart I/Os</td>
<td>The number of actual bytes eligible for predicate offload.</td>
</tr>
<tr>
<td>cell physical IO bytes processed for IM capacity</td>
<td>The number of bytes read from the columnar cache in memcompress for capacity format.</td>
</tr>
<tr>
<td>cell physical IO bytes processed for IM query</td>
<td>The number of bytes read from the columnar cache in memcompress for query format.</td>
</tr>
<tr>
<td>cell physical IO bytes processed for no memcompress</td>
<td>The number of bytes read from the columnar cache in no memcompress format.</td>
</tr>
</tbody>
</table>
### Statistic

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell physical IO bytes saved by columnar cache</td>
<td>The number of bytes saved by columnar cache; that is, the number of bytes of reading that was avoided.</td>
</tr>
<tr>
<td>cell physical IO bytes saved by storage index</td>
<td>The number of bytes saved by storage index.</td>
</tr>
<tr>
<td>cell physical IO bytes saved during optimized file creation</td>
<td>The number of I/O bytes saved by the database host by offloading the file creation operation to the cells. This statistic shows the benefit of optimized file creation operations.</td>
</tr>
<tr>
<td>cell physical IO bytes saved during optimized RMAN restore</td>
<td>The number of I/O bytes saved by the database host by offloading the RMAN file restore operation to the cells. This statistic shows the benefit of optimized RMAN file restore operations.</td>
</tr>
<tr>
<td>cell physical IO bytes sent directly to DB node to balance CPU usage</td>
<td>The number of I/O bytes sent back to the database server for processing due to high storage server CPU usage.</td>
</tr>
<tr>
<td>cell physical IO interconnect bytes</td>
<td>The number of I/O bytes exchanged over the interconnect between the database host and the cells.</td>
</tr>
<tr>
<td>cell physical IO interconnect bytes returned by smart scan</td>
<td>The number of I/O bytes that are returned by the cell for Smart Scan operations. It does not include bytes for other database I/O.</td>
</tr>
</tbody>
</table>

The following table describes database wait events that are useful for monitoring smart I/O operations. The wait events are visible in various dynamic performance views, including `V$SESSION`, `V$SYSTEM_EVENT` and `V$SESSION_EVENT`, and may be displayed in the Wait Event sections of the AWR report.

<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell external table smart scan</td>
<td>This wait event appears when the database is waiting for an external table scan on a cell. The cell hash number in the P1 column in V$SESSION can help to identify a slow cell compared to the rest of the cells.</td>
</tr>
<tr>
<td>cell smart file creation</td>
<td>This wait event appears when the database is waiting for the completion of a file creation on a cell. The cell hash number in the P1 column in V$SESSION can help to identify a slow cell compared to the rest of the cells.</td>
</tr>
<tr>
<td>cell smart incremental backup</td>
<td>This wait event appears when the database is waiting for the completion of an incremental backup on a cell. The cell hash number in the P1 column in V$SESSION can help to identify a slow cell compared to the rest of the cells.</td>
</tr>
<tr>
<td>Wait Event</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>cell smart index scan</td>
<td>This wait event appears when the database is waiting for an index fast full scan. The cell hash number in the P1 column in V$SESSION can help to identify a slow cell when compared to the rest of the cells.</td>
</tr>
<tr>
<td>cell smart index scan: db timezone upgrade</td>
<td>The wait event appears when the cells are unable to offload because a database timezone upgrade is in progress.</td>
</tr>
<tr>
<td>cell smart index scan: disabled by user</td>
<td>The wait event appears when the cells are unable to offload due to a user setting.</td>
</tr>
<tr>
<td>cell smart index scan: pass through</td>
<td>The wait event appears when the cells are unable to offload the Smart Scan.</td>
</tr>
<tr>
<td>cell smart restore from backup</td>
<td>This wait event appears when the database is waiting for the completion of a file initialization for restore from backup on a cell. The cell hash number in the P1 column in V$SESSION can help to identify a slow cell when compared to the rest of the cells.</td>
</tr>
<tr>
<td>cell smart table scan</td>
<td>This wait event appears when the database is waiting for smart scans to complete on a cell. The cell hash number in the P1 column in V$SESSION can help to identify a slow cell when compared to the rest of the cells.</td>
</tr>
<tr>
<td>cell smart table scan: db timezone upgrade</td>
<td>The wait event appears when the cells are unable to offload because a database timezone upgrade is in progress.</td>
</tr>
<tr>
<td>cell smart table scan: disabled by user</td>
<td>The wait event appears when the cells are unable to offload due to a user setting.</td>
</tr>
<tr>
<td>cell smart table scan: pass through</td>
<td>The wait event appears when the cells are unable to offload the Smart Scan.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.

### 7.3.5.3 Monitoring Smart I/O Using SQL Monitor

In addition to dynamic performance views and wait events, Oracle Database provides SQL monitor, which enables you to monitor the execution of individual SQL statements.

The SQL monitor report includes detailed statistics for the row sources, which includes additional information for smart I/O operations. The row source statistics can be viewed in the Enterprise Manager active SQL monitor report by clicking on the binoculars in the row source. For example:
The following table describes various row source statistics provided by SQL monitor, which are useful for monitoring smart I/O.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filtered bytes</td>
<td>The number of bytes returned by the cell.</td>
</tr>
<tr>
<td>Cell passthru I/O bytes</td>
<td>The number of bytes that are not offloaded and sent back to the database for processing.</td>
</tr>
<tr>
<td>Cell passthru I/O bytes due to quarantine</td>
<td>The number of bytes that are not offloaded and sent back to the database for processing due to a quarantine on the cell.</td>
</tr>
<tr>
<td>Eligible bytes for smart I/O</td>
<td>The number of actual bytes eligible for predicate offload. For example, when using columnar cache, this is the size of columnar cache instead of the on-disk size.</td>
</tr>
<tr>
<td>SI saved bytes</td>
<td>The number of bytes saved by storage index; that is, the number of bytes that did not have to be read.</td>
</tr>
<tr>
<td>Columnar cache saved bytes</td>
<td>The number of bytes saved by columnar cache; that is, the number of bytes that did not have to be read.</td>
</tr>
<tr>
<td>Partial flash cache and disk bytes</td>
<td>The number of bytes that read from both Exadata Smart Flash Cache and disk.</td>
</tr>
<tr>
<td>Flash cache bytes</td>
<td>The number of bytes read from Exadata Smart Flash Cache.</td>
</tr>
<tr>
<td>Statistic</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>IM Capacity bytes</td>
<td>The number of bytes read from the columnar cache in memcompress for capacity format.</td>
</tr>
<tr>
<td>IM Query bytes</td>
<td>The number of bytes read from the columnar cache in memcompress for query format.</td>
</tr>
<tr>
<td>No memcompress bytes</td>
<td>The number of bytes read from the columnar cache in no memcompress format.</td>
</tr>
<tr>
<td>Bytes added to storage index</td>
<td>The number of bytes added to the storage index during a Smart Scan. This is an indication that the storage index is being built.</td>
</tr>
<tr>
<td>cell IORM IO requests on flash</td>
<td>The number of physical I/O requests to flash storage.</td>
</tr>
<tr>
<td>cell IORM wait time on flash (us)</td>
<td>The amount of time (in microseconds) IORM queued the flash request. The cell IORM wait time on flash (us)/cell IORM IO requests on flash gives an indication of how much time, on average, is spent in the IORM queue.</td>
</tr>
<tr>
<td>cell IORM IO requests on disk</td>
<td>The number of physical I/O requests to disk storage.</td>
</tr>
<tr>
<td>cell IORM wait time on disk (us)</td>
<td>The amount of time (in microseconds) IORM queued the disk request. The cell IORM wait time on disk (us)/cell IORM IO requests on disk gives an indication of how much time, on average, is spent in the IORM queue.</td>
</tr>
<tr>
<td>Block IO bytes</td>
<td>The number of bytes in block I/O mode.</td>
</tr>
<tr>
<td>Slow metadata bytes</td>
<td>The size of the query metadata sent from the database compute node to a cell.</td>
</tr>
<tr>
<td>Metadata bytes</td>
<td></td>
</tr>
<tr>
<td>Eligible bytes</td>
<td>The number of bytes on-disk eligible for predicate offload.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic is subject to the version of Oracle Database being used.

**Related Topics**

- Real-Time SQL Monitoring and Real-Time Database Operations
- REPORT_SQL_MONITOR Function

### 7.3.5.4 Monitoring Smart I/O Using SQL Explain Plan

The SQL EXPLAIN PLAN command displays information about smart I/O optimizations in the SQL execution plan.

You can use the EXPLAIN PLAN command to identify parts of a SQL query that can be offloaded to the storage server. The database parameter `CELL_OFFLOAD_PLAN_DISPLAY` must be set to AUTO or ALWAYS for the EXPLAIN PLAN command to display the smart I/O optimizations in the SQL execution plan.
Example 7-4 Monitoring Smart I/O Using SQL Explain Plan

This example shows how to use the `EXPLAIN PLAN` command to display the smart I/O optimizations in the SQL execution plan.

In the query plan, the `TABLE ACCESS STORAGE FULL` operation indicates that the corresponding full table scan is offloaded to the storage server. The predicate information further describes the query predicates that are offloaded to the storage server.

```
SQL> ALTER SESSION SET CELL_OFFLOAD_PLAN_DISPLAY = ALWAYS;
Session altered.

SQL> EXPLAIN PLAN FOR
    2   SELECT t.prod_id, v.exp1, t2_prod_id, t2_amount_sold
    3       FROM sales t, v1 v
    4       WHERE t.prod_id = v.prod_id AND t.cust_id = v.cust_id
    5           AND t.prod_id != 45
    6               AND v.amount_sold * v.quantity_sold > 10000;
```

Explained.

```
SQL> SELECT PLAN_TABLE_OUTPUT FROM TABLE(DBMS_XPLAN.DISPLAY());
```

```
PLAN_TABLE_OUTPUT
------------------------------------------------------------------------
<table>
<thead>
<tr>
<th>Id</th>
<th>Operation</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SELECT STATEMENT</td>
<td></td>
</tr>
<tr>
<td>*  1</td>
<td>HASH JOIN</td>
<td></td>
</tr>
<tr>
<td>*  2</td>
<td>HASH JOIN</td>
<td></td>
</tr>
<tr>
<td>*  3</td>
<td>TABLE ACCESS STORAGE FULL</td>
<td>SALES</td>
</tr>
<tr>
<td>*  4</td>
<td>TABLE ACCESS STORAGE FULL</td>
<td>SALES</td>
</tr>
<tr>
<td>*  5</td>
<td>TABLE ACCESS STORAGE FULL</td>
<td>SALES</td>
</tr>
</tbody>
</table>
------------------------------------------------------------------------
Predicate Information (identified by operation id):
------------------------------------------------------------------------
1 - access("T"."CUST_ID"="T2"."CUST_ID" AND
    "T1"."PROD_ID"="T2"."PROD_ID" AND
    "T1"."CUST_ID"="T2"."CUST_ID")
2 - access("T"."PROD_ID"="T1"."PROD_ID")
3 - storage("T1"."PROD_ID"<200 AND
    "T1"."AMOUNT_SOLD"*"T1"."QUANTITY_SOLD">10000 AND
    "T1"."PROD_ID"<>45)
    filter("T1"."PROD_ID"<200 AND
        "T1"."AMOUNT_SOLD"*"T1"."QUANTITY_SOLD">10000 AND
        "T1"."PROD_ID"<>45)
4 - storage("T"."PROD_ID"<200 AND "T"."PROD_ID"<>45)
    filter("T"."PROD_ID"<200 AND "T"."PROD_ID"<>45)
------------------------------------------------------------------------
7.3.5.5 Monitoring Smart I/O Using Exadata Metrics

Smart I/O metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=SMARTIO.

For cumulative metrics, the metric value for a specific time period can be determined by subtracting values from different collectionTime periods.

Example 7-5  Displaying Smart I/O Metric Definitions

This example shows how to display the smart I/O metric definitions that are available in the Oracle Exadata System Software.

CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = SMARTIO

SIO_IO_EL_OF  "Cumulative number of megabytes eligible for smart IO offload"
SIO_IO_EL_OF_SEC  "Number of megabytes per second eligible for smart IO offload"
SIO_IO_OF_RE  "Cumulative number of interconnect megabytes returned by smart IO"
SIO_IO_OF_RE_SEC  "Number of interconnect megabytes per second returned by smart IO"
SIO_IO_PA_TH  "Cumulative number of megabytes of passthru I/Os by smart IO"
SIO_IO_PA_TH_SEC  "Number of megabytes per second of passthru I/Os by smart IO"
SIO_IO_RD_FC  "Cumulative number of megabytes read from flash cache by smart IO"
SIO_IO_RD_FC_HD  "Cumulative number of megabytes read from both flash cache and hard disk by smart IO"
SIO_IO_RD_FC_HD_SEC  "Number of megabytes per second read from both flash cache and hard disk by smart IO"
SIO_IO_RD_FC_SEC  "Number of megabytes per second read from flash cache by smart IO"
SIO_IO_RD_HD  "Cumulative number of megabytes read from hard disk by smart IO"
SIO_IO_RD_HD_SEC  "Number of megabytes per second read from hard disk by smart IO"
SIO_IO_RD_RQ_FC  "Cumulative number of read IO requests from flash cache by smart IO"
SIO_IO_RD_RQ_FC_HD  "Cumulative number of read IO requests from both flash cache and hard disk by smart IO"
SIO_IO_RD_RQ_FC_HD_SEC  "Number of read IO requests per second from both flash cache and hard disk by smart IO"
SIO_IO_RD_RQ_FC_SEC  "Number of read IO requests per second from flash cache by smart IO"
SIO_IO_RD_RQ_HD  "Cumulative number of read IO requests from hard disk by smart IO"
SIO_IO_RD_RQ_HD_SEC  "Number of read IO requests per second from hard disk by smart IO"
Columnar cache metrics are identified as having name like 'FC_COL.*'

Example 7-6  Displaying Columnar Cache Metric Definitions

This example shows how to display the columnar cache metric definitions that are available in the Oracle Exadata System Software.

`CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE NAME LIKE 'FC_COL.*'
FC_COL_BYKEEP_USED           "Number of megabytes used for keep objects in Columnar FlashCache"
FC_COL_BY_USED               "Number of megabytes used in Columnar FlashCache"
FC_COL_IO_BYKEEP_R           "Number of megabytes read from Columnar FlashCache for keep objects"
FC_COL_IO_BYKEEP_R_SEC      "Number of megabytes per second that were read from Columnar FlashCache"
FC_COL_IO_BY_R_ELIGIBLE      "Number of megabytes that were eligible to read from Columnar FlashCache"
FC_COL_IO_BY_R_ELIGIBLE_SEC "Number of megabytes per second eligible to read from Columnar FlashCache"
FC_COL_IO_BY_SAVED           "Number of megabytes saved by reads from Columnar FlashCache"
FC_COL_IO_BY_SAVED_SEC      "Number of megabytes saved per second by reads from Columnar FlashCache"
FC_COL_IO_BY_W_POPULATE     "Number of megabytes that are population writes into Columnar FlashCache due to read miss"
7.3.5.6 What to Look For When Monitoring Smart I/O

Smart I/O operations typically occur when a full table scan or index fast full scan is done on a row source. If smart I/O operations do not work as expected, then users tend to experience noticeable increases in query elapsed times. In some cases, the database shows increased cell smart table scan wait times. However, when wait events such as cell multiblock physical read or direct path read are present instead of cell smart table scan, this is an indicator that smart I/O operations are not being performed.

The following are symptoms and reasons that explain why smart I/O operations might not perform as expected:

- Direct reads are a prerequisite for Smart Scan, and without direct reads Smart Scan cannot occur. The cell multiblock physical read wait event occurs when blocks are read into the buffer cache. A common reason for reading into the buffer cache, rather than using direct reads, is the size of the segment. For small segments, the optimizer may choose buffer cache reads instead of direct reads.

- The direct path read wait event occurs when direct reads are performed, but predicates are not offloaded to the storage servers. This may occur when there is a resource shortage on the storage servers. For example, there might be a memory shortage caused by a large number of concurrent parallel queries on the system.

Such resource shortages are typically evident in the ExaWatcher data. In particular, you can review statistics such as Number of low memory threshold failures and Number of no memory threshold failures in cellsrvstat. ExaWatcher also includes cellmem collection, which shows how memory is consumed on the storage servers, and is visible in the ExaWatcher charts produced using GetExaWatcherResults.sh.

To address this issue, you can review the use of parallel query and potentially reduce the number of active parallel query servers.
Predicate offload is not possible when there is uncommitted data. This typically becomes an issue if large batch operations are modifying the data, and you attempt to run a large query against a large amount of uncommitted data.

When Smart Scan encounters uncommitted data, predicate filtering cannot be offloaded to the storage servers, and additional data must be transported back to the database servers, which appears as increased bytes returned by the Smart Scan. Extra processing is also required on the database server to construct a read-consistent copy of the data, which is manifested in the following ways:

- In the best-case scenario, additional buffer gets or session logical reads are required to construct a read-consistent copy of the data.
- If the undo buffers reside in another database instance, then Oracle RAC-related wait events may also be observed. The Oracle RAC-related wait events are prefixed by `gc`.
- If the undo blocks do not reside in the buffer cache of any database instance, then additional `cell single block physical read` waits are observed in conjunction with the single block I/O that is required for read-consistency. The additional I/O can significantly impact the performance of the operation.

When the storage server CPU utilization is high, the storage server sends data back to the database for processing, rather than consuming even more storage server CPU to perform predicate evaluation. This is known as 'reverse offload'.

When this occurs, it is evident in the Reverse Offload column in the Smart IO section of the AWR report, and the database statistic `cell physical IO bytes sent directly to DB node to balance CPU usage`.

The high CPU utilization on the storage servers may be due to the type of predicates being offloaded. For example, case-insensitive searches, or the use of `REGEXP_LIKE` utilize more CPU than simpler predicates.

Increased storage server CPU and I/O load may also stem from SQL execution plan changes in the database, in which case, reviewing the execution plans and tuning the affected SQL statements may help to resolve the issue.

When the storage server is unable to perform predicate evaluation, it will send the data back to the database for processing. This is also known as passthough (or 'passthru'). The following all indicate that passthrough is occurring:

- A large value in the Passthru column in the Smart IO section of the AWR report, when compared to eligible bytes.
- A large value in the database statistic `cell num bytes in passthru during predicate offload, when compared to cell physical IO bytes eligible for smart IO`.
- A large `Cell passthru IO bytes` value in the SQL monitor row source statistic, compared to the Eligible bytes for Smart IO value.

Possible causes include:

- Quarantines — To confirm, review the database statistic `cell num bytes in passthru due to quarantine, or the SQL monitor row source statistic Cell passthru IO bytes due to quarantine`.
- Database timezone upgrade — Smart Scan is disabled when a database timezone upgrade is ongoing. Review the database statistic `cell num smart IO sessions using passthru mode due to timezone` from the Global Activity Statistics or Instance Activity Statistics section, or the Passthru
Reasons in the Smart IO section of the AWR report. Depending on the database release, you may also observe the cell smart table scan: db timezone upgrade or cell smart index scan: db timezone upgrade wait event.

– User setting — A user or application may set cell_offload_processing=false, which disables smart scans. To confirm, review the database statistic cell num smart IO sessions using passthru mode due to user from the Global Activity Statistics or Instance Activity Statistics section, or the Passthru Reasons in the Smart IO section of the AWR report. Depending on the database release, you may also observe the cell smart table scan: disabled by user or cell smart index scan: disabled user wait event.

– Operation cannot be offloaded — There are other reasons why the storage servers may be unable to perform predicate offload. Instances of this occurrence would be visible in the database statistic cell num smart IO sessions using passthru mode due to cellsrv, or in wait events cell smart table scan: passthrough or cell smart index scan: passthrough. The following section describes the reasons in detail.

Operation Not Being Offloaded

A smart I/O operation cannot be offloaded to the Exadata storage servers in the following cases:

• A scan is performed on a clustered table
• A scan is performed on an index-organized table
• A fast full scan is performed on compressed indexes
• A fast full scan is performed on reverse key indexes
• The table has row dependencies enabled or the rowscn is being fetched
• The optimizer wants the scan to return rows in ROWID order
• The command CREATE INDEX using nosort
• A LOB or LONG column is being selected or queried.
• A SELECT ... VERSIONS query is done on a table
• A query that has more than 255 columns referenced, and the heap table is uncompressed, or Basic or OLTP compressed. However, such queries on tables compressed using Exadata Hybrid Columnar Compression are offloaded.

• The tablespace is encrypted, and the CELL_OFFLOAD_DECRYPTION parameter is set to FALSE. In order for the Oracle Exadata System Software to perform decryption, Oracle Database needs to send the decryption key to the storage server. This feature is typically disabled, if there are security concerns about keys being shipped across the network to the storage server
• The tablespace is not completely stored on Oracle Exadata Storage Server.
• The predicate evaluation is on a virtual column.

• Although offloading is supported for most SQL operators and functions, Oracle Exadata System Software does not support offloading for some SQL operators and functions. The dynamic performance view V$SQLFN_METADATA includes information about whether offloading is supported for a SQL operator or function. If the OFFLOADABLE column contains YES, then offloading is supported for the
corresponding operator or function. NO indicates that offloading is not supported for the corresponding operator or function.

Storage Index Not Performing as Expected

To monitor storage index performance, monitor the Storage Index column in the Smart IO section of the AWR report, or the database statistic cell physical IO bytes saved by storage index, or the SQL monitor row source statistic SI saved bytes.

The storage index is maintained in cellsrv memory and is rebuilt every time cellsrv starts. Consequently, storage index cannot benefit operations immediately after cellsrv starts.

Also, for segments in an unencrypted tablespace, the storage index is maintained when DML occurs. However, for segments in an encrypted tablespace, DML invalidates the portion of the storage index that is associated with each changed data chunk (1 MB). The invalidated chunks are rebuilt during the next scan of the segment. However, the overall efficiency of the storage index is not optimal while portions are invalid.

You may see reduced storage index savings if the storage index is not yet built (or rebuilt). If this is the case, you will see increases in the database statistic cell physical IO bytes added to storage index, or the SQL monitor row source statistic Bytes added to storage index.

Columnar Cache Not Performing as Expected

Like storage index, if the columnar cache is not yet built (or rebuilt), you may see reduced savings associated with columnar cache. To monitor columnar cache performance, monitor the Columnar Cache sections of the AWR report, or the database statistic cell physical IO bytes saved by columnar cache, or the SQL monitor row source statistic Columnar cache saved bytes.

Similar to storage index, the columnar cache is rebuilt every time cellsrv starts. Consequently, the columnar cache cannot benefit operations immediately after cellsrv starts.

The columnar cache is automatically populated and maintained by the storage server when a Smart Scan is performed. For Smart Scan operations on uncompressed segments, segments compressed using OLTP compression, and segments compressed using Exadata Hybrid Columnar Compression, data is automatically converted into the columnar cache format (no memcompress) as part of the Smart Scan.

However, if you are using Oracle Database In-Memory, data is rewritten into Oracle Database In-Memory columnar format (memcompress for query or memcompress for capacity) by using a background process. Consequently, operations on that data do not benefit from the Oracle Database In-Memory optimizations until the cache is repopulated. Information about the population jobs is available in the Columnar Cache Population sections of the AWR report.

If you are reading less from columnar cache, then it will be evident by lower values in the database statistics: cell physical IO bytes processed for IM Query, cell physical IO bytes processed for IM Capacity, or cell physical IO bytes processed for no memcompress. The equivalent SQL monitor row source statistics are: IM Query bytes, IM Capacity bytes, and No memcompress bytes.
7.3.6 Monitoring I/O Resource Management (IORM)

Exadata I/O Resource Management (IORM) enables multiple databases and pluggable databases (PDBs) to share the same storage while ensuring that I/O resources are allocated appropriately across the various databases that share the system. This prevents one database from utilizing the entire I/O bandwidth and degrading the performance of the other databases. IORM also works in conjunction with Oracle Database Resource Manager to manage I/O resources across multiple consumer groups within a single database.

- Monitoring I/O Resource Management (IORM) Using AWR
- Monitoring I/O Resource Management (IORM) Using Database Statistics
- Monitoring I/O Resource Management (IORM) Using Exadata Metrics
- What to Look For When Monitoring I/O Resource Management (IORM)

7.3.6.1 Monitoring I/O Resource Management (IORM) Using AWR

The AWR report includes information about the top databases running on the Exadata system. By using this information, you can monitor system-wide I/O resource consumption regardless of the database that hosts the AWR report. Additionally, the AWR report includes per-cell statistics, which enables you to easily determine if there is any skew in the I/O resource usage across the cells.

The AWR report includes:

- Top Databases by IO Requests — provides statistics for the databases with the greatest number of I/O requests across all cells. It also breaks down the requests by device type (flash or disk) and by size (small or large).

- Top Databases by Requests - Details — provides additional statistics regarding the I/O requests. Specifically, it shows the average latency and the average queue time for I/O requests. The queue time is the amount of time that an I/O requests spends in relevant I/O queue. Large queue times indicate that IORM is throttling the I/O.

- Top Databases by IO Requests per Cell — provides statistics for the databases with the greatest number of I/O requests on each cell. It also breaks down the requests by device type (flash or disk) and by size (small or large). By using this information, you can easily determine if any cell behaves differently when servicing the I/O requests from the database(s).

- Top Databases by IO Requests per Cell - Details — provides additional statistics regarding the I/O requests on each cell. Specifically, it shows the average latency and the average queue time for I/O requests on each cell. By using this information, you can easily determine if IORM behaves differently on any cell.

- Top Databases by IO Throughput — provides statistics for the databases with the greatest I/O throughput across all cells. It also breaks down the throughput by device type (flash or disk) and by request size (small or large).

- Top Databases by Throughput per Cell — provides statistics for the databases with the greatest I/O throughput on each cell. It also breaks down the throughput by device type (flash or disk) and by request size (small or large). By using this information, you can easily determine if any cell behaves differently when servicing the I/O from the database(s).
The AWR report may not be show information about all of the databases running on the Exadata system. Rather, it is designed to focus on the databases that are responsible for the majority of I/O on the storage servers. If more detailed information is required, then IORM cell metrics should be used.

For multitenant container databases (CDBs), the statistics in the AWR report include all of the I/Os associated with the database, including all of the pluggable databases (PDBs).

The following example shows the Top Databases by IO Requests and Top Databases by Requests - Details sections in the AWR report. The Top Databases by IO Requests section shows that the DB1 database generates 46% of the I/O requests captured in AWR. It also shows that approximately 10% of the I/O requests associated with the DB1 database are I/O requests to disk devices (51,854,884 disk I/O requests of the 564,616,086 total I/O requests). The Top Databases by Requests - Details section mostly shows minimal IORM wait time (queue time < 1 millisecond) across all databases. The exception is large I/O requests to flash for the DB1 database, which each have an average queue time of 9.12 milliseconds.

![AWR Report: Top Databases by IO Requests](image)

**Figure 7-19 AWR Report: Top Databases by IO Requests**

<table>
<thead>
<tr>
<th></th>
<th>DB1 (*)</th>
<th>DB2</th>
<th>OTHER</th>
<th>ASM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Name</td>
<td>3456789012</td>
<td>987654321</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DBID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%Captured</td>
<td>46.6%</td>
<td>27.5%</td>
<td>25.8%</td>
<td>0.58%</td>
</tr>
<tr>
<td>Total I/O</td>
<td>564,616,086</td>
<td>338,628,595</td>
<td>317,150,736</td>
<td>7,067,228</td>
</tr>
<tr>
<td>Requests</td>
<td>52,284,112</td>
<td>31,357,403</td>
<td>300,520,056</td>
<td>645,432</td>
</tr>
<tr>
<td>per Sec</td>
<td>512,761,202</td>
<td>331,392,170</td>
<td>16,830,680</td>
<td>3,223,716</td>
</tr>
<tr>
<td>Flash I/O</td>
<td>51,854,884</td>
<td>7,236,429</td>
<td>1,903,37</td>
<td>3,843,512</td>
</tr>
<tr>
<td>Disk I/O</td>
<td>44,861,630</td>
<td>20,564,529</td>
<td>17,503,205</td>
<td>429,701</td>
</tr>
<tr>
<td>MB</td>
<td>27,244,522</td>
<td>1,475,605</td>
<td>9,492,003</td>
<td>126,257</td>
</tr>
<tr>
<td>Throughput</td>
<td>24,422,862</td>
<td>21,722,104</td>
<td>8,011,202</td>
<td>303,443</td>
</tr>
</tbody>
</table>

**Top Databases By Requests - Details**

- Request details for the top databases by I/O requests

<table>
<thead>
<tr>
<th></th>
<th>Small Requests</th>
<th>Large Requests</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Name</td>
<td>DBID</td>
<td>Latency</td>
</tr>
<tr>
<td>DB1 (*)</td>
<td>52,284,112</td>
<td>210.10</td>
</tr>
<tr>
<td>DB2</td>
<td>31,357,403</td>
<td>3,586.87</td>
</tr>
<tr>
<td>OTHER</td>
<td>29,368,532</td>
<td>5,6726.38</td>
</tr>
<tr>
<td>ASM</td>
<td>654.43</td>
<td>449.17</td>
</tr>
</tbody>
</table>

7.3.6.2 Monitoring I/O Resource Management (IORM) Using Database Statistics

The following table describes database statistics that are useful for monitoring I/O Resource Management (IORM). The statistics are available in various dynamic performance views, including $SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session total flash IO requests</td>
<td>Total number of physical I/O requests on flash.</td>
</tr>
<tr>
<td>Session IORM flash wait time</td>
<td>Total amount of IORM wait time for flash I/O requests (in microseconds). To determine the average IORM wait time for flash I/O requests, divide Session IORM flash wait time by Session total flash IO requests.</td>
</tr>
<tr>
<td>PDB total disk IO requests</td>
<td>Total number of physical I/O requests on disk.</td>
</tr>
<tr>
<td>PDB IORM disk wait time</td>
<td>Total amount of IORM wait time for disk I/O requests (in microseconds). To determine the average IORM wait time for disk I/O requests, divide PDB IORM disk wait time by PDB total disk IO requests.</td>
</tr>
</tbody>
</table>

The availability of a specific statistic is subject to the version of Oracle Database being used.

### 7.3.6.3 Monitoring I/O Resource Management (IORM) Using Exadata Metrics

You can monitor I/O Resource Management (IORM) by using Oracle Exadata System Software metrics.

IORM uses the database unique name, not the database identifier, to collect statistics and display output. Starting with Oracle Exadata System Software release 19.1.0, if you configured ASM-scoped security for the Oracle Automatic Storage Management (Oracle ASM) cluster used by the database, then the database name is prefixed with the Oracle ASM cluster name.

- **Monitoring IORM with Database Metrics**
  Database metrics provide information about the I/O load from each database listed in the IORM interdatabase plan.

- **Monitoring IORM with PDB Metrics**
  Pluggable Database (PDB) metrics provide information about the I/O load from each PDB hosted by a container database (CDB) listed in the IORM interdatabase plan.

- **Monitoring IORM with Consumer Group Metrics**
  Consumer group metrics provide information about the I/O load from each consumer group specified in a database resource plan.

- **Monitoring IORM with Category Metrics**
  Category metrics provide information about the I/O load from each category specified in the current IORM category plan.

- **Monitoring IORM Utilization**
  You can use metrics to monitor IORM utilization.
7.3.6.3.1 Monitoring IORM with Database Metrics

Database metrics provide information about the I/O load from each database listed in the IORM interdatabase plan.

Database metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=IORM_DATABASE.

Example 7-7  Displaying Category Metric Definitions

This example shows how to display the database metric definitions that are available in the Oracle Exadata System Software.

CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = IORM_DATABASE
DB_FC_BYALLOCATED  "Number of megabytes allocated in
flash cache for this database"
DB_FC_IO_BY_SEC    "Number of megabytes of I/O per second
for this database to flash cache"
DB_FC_IO_RQ        "Number of IO requests issued by a
database to flash cache"
DB_FC_IO_RQ_LG     "Number of large IO requests issued by
a database to flash cache"
DB_FC_IO_RQ_LG_SEC "Number of large IO requests issued by
a database to flash cache per second"
DB_FC_IO_RQ_SEC    "Number of IO requests issued by a
database to flash cache per second"
DB_FC_IO_RQ_SM     "Number of small IO requests issued by
a database to flash cache"
DB_FC_IO_RQ_SM_SEC "Number of small IO requests issued by
a database to flash cache per second"
DB_FD_IO_BY_SEC    "Number of megabytes of I/O per second
for this database to flash disks"
DB_FD_IO_LOAD      "Average I/O load from this database
for flash disks"
DB_FD_IO_RQ_LG     "Number of large IO requests issued by
a database to flash disks"
DB_FD_IO_RQ_LG_SEC "Number of large IO requests issued by
a database to flash disks per second"
DB_FD_IO_RQ_SM     "Number of small IO requests issued by
a database to flash disks"
DB_FD_IO_RQ_SM_SEC "Number of small IO requests issued by
a database to flash disks per second"
DB_FD_IO_TM        "The cumulative latency of reading or
writing blocks by a database from flash disks"
DB_FD_IO_TM_LG     "The cumulative latency of reading or
writing large blocks by a database from flash disks"
DB_FD_IO_TM_LG_RQ  "The rate which is the average latency
of reading or writing large blocks per request by a database from flash disks"
DB_FD_IO_TM_RQ     "The rate which is the average latency
of reading or writing blocks per request by a database from flash disks"
DB_FD_IO_TM_SM     "The cumulative latency of reading or
writing small blocks by a database from flash disks"
DB_FD_IO_TM_SM_RQ  "The rate which is the average latency
of reading or writing small blocks per request by a database from flash disks"
of reading or writing small blocks per request by a database from flash disks"

DB_FD_IO_UTIL "Percentage of flash resources utilized by this database"
DB_FD_IO_UTIL_LG "Percentage of flash resources utilized by large requests from this database"
DB_FD_IO_UTIL_SM "Percentage of flash resources utilized by small requests from this database"
DB_FD_IO_WT_LG "IORM wait time for large IO requests issued to flash disks by a database"
DB_FD_IO_WT_LG_RQ "Average IORM wait time per request for large IO requests issued to flash disks by a database"
DB_FD_IO_WT_SM "IORM wait time for small IO requests issued to flash disks by a database"
DB_FD_IO_WT_SM_RQ "Average IORM wait time per request for small IO requests issued to flash disks by a database"

DB_FL_IO_BY "The number of MB written to the Flash Log"
DB_FL_IO_BY_SEC "The number of MB written per second to the Flash Log"
DB_FL_IO_RQ "The number of I/O requests issued to the Flash Log"
DB_FL_IO_RQ_SEC "The number of I/O requests per second issued to the Flash Log"

DB_IO_BY_SEC "Number of megabytes of I/O per second for this database to hard disks"
DB_IO_LOAD "Average I/O load from this database for hard disks"
DB_IO_RQ_LG "Number of large IO requests issued by a database to hard disks"
DB_IO_RQ_LG_SEC "Number of large IO requests issued by a database to hard disks per second"
DB_IO_RQ_SM "Number of small IO requests issued by a database to hard disks"
DB_IO_RQ_SM_SEC "Number of small IO requests issued by a database to hard disks per second"

DB_IO_TM_LG "The cumulative latency of reading or writing large blocks by a database from hard disks"
DB_IO_TM_LG_RQ "The rate which is the average latency of reading or writing large blocks per request by a database from hard disks"
DB_IO_TM_SM "The cumulative latency of reading or writing small blocks by a database from hard disks"
DB_IO_TM_SM_RQ "The rate which is the average latency of reading or writing small blocks per request by a database from hard disks"

DB_IO_UTIL_LG "Percentage of disk resources utilized by large requests from this database"
DB_IO_UTIL_SM "Percentage of disk resources utilized by small requests from this database"
DB_IO_WT_LG "IORM wait time for large IO requests issued to hard disks by a database"
DB_IO_WT_LG_RQ "Average IORM wait time per request for large IO requests issued to hard disks by a database"
DB_IO_WT_SM "IORM wait time for small IO requests issued to hard disks by a database"
issued to hard disks by a database"

DB_IO_WT_SM_RQ          "Average IORM wait time per request for small IO requests issued to hard disks by a database"
DB_PC_BY_ALLOCATED      "Number of megabytes allocated in PMEM cache for this database"

Note the following additional details:

- The database for the metric is specified by the metricObjectName attribute of the METRICCURRENT and METRICHISTORY objects.

- For metrics that relate to I/O load (for example, DB_FD_IO_LOAD and DB_IO_LOAD), see the additional information relating to CD_IO_LOAD.

- Starting with Oracle Exadata System Software release 19.1.0, if you configured ASM-scoped security for the Oracle Automatic Storage Management (Oracle ASM) cluster used by the database, then the database name is prefixed with the Oracle ASM cluster name.

- For cumulative metrics, the metric value for a specific time period can be determined by subtracting values from different collectionTime periods.

- For rate metrics, the time period for the metric value is over the previous minute.

- In the metric descriptions, small I/O requests are less than or equal to 128 KB, and large I/O requests are larger than 128 KB.

- All database cumulative metrics are reset to zero whenever a category, IORM, or any database resource plan is modified.

- To list the database metric history for an inter-database plan, use the following CellCLI command:

  CellCLI> LIST METRICHISTORY WHERE objectType = 'IORM_DATABASE' AND metricValue != 0 ATTRIBUTES name, metricObjectName, metricValue, collectionTime

- For multitenant container databases (CDBs), the database metric observations include all of the I/Os associated with the database, including all of the associated pluggable databases (PDBs). For example, the value for DB_FC_IO_BY_SEC includes the sum of the PDB_FC_IO_BY_SEC values for all of the PDBs hosted by the CDB.

- Observations for Oracle ASM and all other databases not listed in the interdatabase plan are grouped together using _OTHER_DATABASE_ as the metricObjectName value.

7.3.6.3.2 Monitoring IORM with PDB Metrics

Pluggable Database (PDB) metrics provide information about the I/O load from each PDB hosted by a container database (CDB) listed in the IORM interdatabase plan.

PDB metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=IORM_PLUGGABLE_DATABASE.
Example 7-8  Displaying PDB Metric Definitions

This example shows how to display the PDB metric definitions that are available in the Oracle Exadata System Software.

CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = IORM_PLUGGABLE_DATABASE

<table>
<thead>
<tr>
<th>Metric Definition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB_FC_BY_ALLOCATED</td>
<td>Number of megabytes allocated in flash cache for this pluggable database</td>
</tr>
<tr>
<td>PDB_FC_IO_BY_SEC</td>
<td>Number of megabytes of I/O per second for this pluggable database to flash cache</td>
</tr>
<tr>
<td>PDB_FC_IO_RQ</td>
<td>Number of I/O requests issued by this pluggable database to flash cache</td>
</tr>
<tr>
<td>PDB_FC_IO_RQ_LG</td>
<td>Number of large I/O requests issued by this pluggable database to flash cache</td>
</tr>
<tr>
<td>PDB_FC_IO_RQ_LG_SEC</td>
<td>Number of large I/O requests issued by this pluggable database to flash cache per second</td>
</tr>
<tr>
<td>PDB_FC_IO_RQ_SM</td>
<td>Number of small I/O requests issued by this pluggable database to flash cache</td>
</tr>
<tr>
<td>PDB_FC_IO_RQ_SM_SEC</td>
<td>Number of small I/O requests issued by this pluggable database to flash cache per second</td>
</tr>
<tr>
<td>PDB_FD_IO_BY_SEC</td>
<td>Number of megabytes of I/O per second for this pluggable database to flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_LOAD</td>
<td>Average I/O load from this pluggable database for flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_RQ_LG</td>
<td>Number of large I/O requests issued by this pluggable database to flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_RQ_LG_SEC</td>
<td>Number of large I/O requests issued by this pluggable database to flash disks per second</td>
</tr>
<tr>
<td>PDB_FD_IO_RQ_SM</td>
<td>Number of small I/O requests issued by this pluggable database to flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_RQ_SM_SEC</td>
<td>Number of small I/O requests issued by this pluggable database to flash disks per second</td>
</tr>
<tr>
<td>PDB_FD_IO_TM</td>
<td>The cumulative latency of reading or writing blocks by this pluggable database from flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_TM_LG</td>
<td>The cumulative latency of reading or writing large blocks by this pluggable database from flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_TM_RQ</td>
<td>The rate which is the average latency of reading or writing large blocks per request by this pluggable database from flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_TM_SM</td>
<td>The cumulative latency of reading or writing small blocks by this pluggable database from flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_TM_SM_RQ</td>
<td>The rate which is the average latency of reading or writing small blocks per request by this pluggable database from flash disks</td>
</tr>
<tr>
<td>PDB_FD_IO_UTIL</td>
<td>Percentage of flash resources utilized by this pluggable database</td>
</tr>
<tr>
<td>PDB_FD_IO_UTIL_LG</td>
<td>Percentage of flash resources utilized by large requests from this pluggable database</td>
</tr>
<tr>
<td>PDB_FD_IO_UTIL_SM</td>
<td>Percentage of flash resources utilized by small requests from this pluggable database</td>
</tr>
</tbody>
</table>
utilized by small requests from this pluggable database"
PDB_FD_IO_WT_SM "IORM wait time for small IO requests issued to flash disks by this pluggable database"
PDB_FD_IO_WT_SM_RQ "Average IORM wait time per request for small IO requests issued to flash disks by this pluggable database"
PDB_IO_BY_SEC "Number of megabytes of I/O per second for this pluggable database to hard disks"
PDB_IO_LOAD "Average I/O load from this pluggable database for hard disks"
PDB_IO_RQ_LG "Number of large IO requests issued by this pluggable database to hard disks"
PDB_IO_RQ_LG_SEC "Number of large IO requests issued by this pluggable database to hard disks per second"
PDB_IO_RQ_SM "Number of small IO requests issued by this pluggable database to hard disks"
PDB_IO_RQ_SM_SEC "Number of small IO requests issued by this pluggable database to hard disks per second"
PDB_IO_TM_LG "The cumulative latency of reading or writing large blocks by this pluggable database from hard disks"
PDB_IO_TM_LG_RQ "The rate which is the average latency of reading or writing large blocks per request by this pluggable database from hard disks"
PDB_IO_TM_SM "The cumulative latency of reading or writing small blocks by this pluggable database from hard disks"
PDB_IO_TM_SM_RQ "The rate which is the average latency of reading or writing small blocks per request by this pluggable database from hard disks"
PDB_IO_UTIL_LG "Percentage of disk resources utilized by large requests from this pluggable database"
PDB_IO_UTIL_SM "Percentage of disk resources utilized by small requests from this pluggable database"
PDB_IO_WT_LG "IORM wait time for large IO requests issued to hard disks by this pluggable database"
PDB_IO_WT_LG_RQ "Average IORM wait time per request for large IO requests issued to hard disks by this pluggable database"
PDB_IO_WT_SM "IORM wait time for small IO requests issued to hard disks by this pluggable database"
PDB_IO_WT_SM_RQ "Average IORM wait time per request for small IO requests issued to hard disks by this pluggable database"
PDB_PC_BY_ALLOCATED "Number of megabytes allocated in PMEM cache for this pluggable database"

Note the following additional details:

- The PDB for the metric is specified by the metricObjectName attribute of the METRICCURRENT and METRICHISTORY objects. The PDB name is a concatenation of the CDB name with the PDB name.
- For metrics that relate to I/O load (for example, PDB_FD_IO_LOAD and PDB_IO_LOAD), see the additional information relating to CD_IO_LOAD.
Starting with Oracle Exadata System Software release 19.1.0, if you configured ASM-scoped security for the Oracle Automatic Storage Management (Oracle ASM) cluster used by the database, then the database name is prefixed with the Oracle ASM cluster name.

For cumulative metrics, the metric value for a specific time period can be determined by subtracting values from different collectionTime periods.

For rate metrics, the time period for the metric value is over the previous minute.

In the metric descriptions, small I/O requests are less than or equal to 128 KB, and large I/O requests are larger than 128 KB.

7.3.6.3.3 Monitoring IORM with Consumer Group Metrics

Consumer group metrics provide information about the I/O load from each consumer group specified in a database resource plan.

Consumer group metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=IORM_CONSUMER_GROUP.

Example 7-9 Displaying Consumer Group Metric Definitions

This example shows how to display the consumer group metric definitions that are available in the Oracle Exadata System Software.

```sql
CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = IORM_CONSUMER_GROUP
CG_FC_IO_BY_SEC  "Number of megabytes of I/O per second for this consumer group to flash cache"
CG_FC_IO_RQ      "Number of IO requests issued by a consumer group to flash cache"
CG_FC_IO_RQ_LG   "Number of large IO requests issued by a consumer group to flash cache"
CG_FC_IO_RQ_SEC  "Number of IO requests issued by a consumer group to flash cache per second"
CG_FC_IO_RQ_SM   "Number of small IO requests issued by a consumer group to flash cache"
CG_FC_IO_RQ_SM_SEC  "Number of small IO requests issued by a consumer group to flash cache per second"
CG_FD_IO_By_SEC  "Number of megabytes of I/O per second for this consumer group to flash disks"
CG_FD_IO_LOAD    "Average I/O load from this consumer group for flash disks"
CG_FD_IO_RQ_LG   "Number of large IO requests issued by a consumer group to flash disks"
CG_FD_IO_RQ_LG_SEC  "Number of large IO requests issued by a consumer group to flash disks per second"
CG_FD_IO_RQ_SM   "Number of small IO requests issued by a consumer group to flash disks"
CG_FD_IO_RQ_SM_SEC  "Number of small IO requests issued by a consumer group to flash disks per second"
CG_FD_IO_TM     "The cumulative latency of reading or writing blocks by a consumer group from flash disks"
CG_FD_IO_TM_LG  "The cumulative latency of reading or writing blocks by a consumer group from flash disks"
```
writing large blocks by a consumer group from flash disks"

CG_FD_IO_TM_LG_RQ  "The rate which is the average latency
of reading or writing large blocks per request by a consumer group from
flash disks"

CG_FD_IO_TM_RQ  "The rate which is the average latency
of reading or writing blocks per request by a consumer group from flash
disks"

CG_FD_IO_TM_SM  "The cumulative latency of reading or
writing small blocks by a consumer group from flash disks"

CG_FD_IO_TM_SM_RQ  "The rate which is the average latency
of reading or writing small blocks per request by a consumer group from
flash disks"

CG_FD_IO_UTIL  "Percentage of flash resources
utilized by this consumer group"

CG_FD_IO_UTIL_LG  "Percentage of flash resources
utilized by large requests from this consumer group"

CG_FD_IO_UTIL_SM  "Percentage of flash resources
utilized by small requests from this consumer group"

CG_FD_IO_WT_LG  "IORM wait time for large IO requests
issued to flashdisks by a consumer group"

CG_FD_IO_WT_LG_RQ  "Average IORM wait time per request
for large IO requests issued to flash disks by a consumer group"

CG_FD_IO_WT_SM  "IORM wait time for small IO requests
issued to flashdisks by a consumer group"

CG_FD_IO_WT_SM_RQ  "Average IORM wait time per request
for small IO requests issued to flash disks by a consumer group"

CG_IO_BY_SEC  "Number of megabytes of I/O per second
for this consumer group to hard disks"

CG_IO_LOAD  "Average I/O load from this consumer
group for hard disks"

CG_IO_RQ_LG  "Number of large IO requests issued by
a consumer group to hard disks"

CG_IO_RQ_LG_SEC  "Number of large IO requests issued by
a consumer group to hard disks per second"

CG_IO_RQ_SM  "Number of small IO requests issued by
a consumer group to hard disks"

CG_IO_RQ_SM_SEC  "Number of small IO requests issued by
a consumer group to hard disks per second"

CG_IO_TM_LG  "The cumulative latency of reading or
writing large blocks by a consumer group from hard disks"

CG_IO_TM_LG_RQ  "The rate which is the average latency
of reading or writing large blocks per request by a consumer group from
hard disks"

CG_IO_UTIL_LG  "Percentage of disk resources utilized
by large requests from this consumer group"

CG_IO_UTIL_SM  "Percentage of disk resources utilized
by small requests from this consumer group"

CG_IO_WT_LG  "IORM wait time for large IO requests
issued to hard disks by a consumer group"

CG_IO_WT_LG_RQ  "Average IORM wait time per request
for large IO requests issued to hard disks by a consumer group"
for large IO requests issued to hard disks by a consumer group"

CG_IO_WT_SM             "IORM wait time for small IO requests issued to hard disks by a consumer group"

CG_IO_WT_SM_RQ          "Average IORM wait time per request for small IO requests issued to hard disks by a consumer group"

Note the following additional details:

- The consumer group and database for the metric are specified by the `metricObjectName` attribute of the `METRICCURRENT` and `METRICHISTORY` objects. The name is formed by the database name followed by a period (.) and the consumer group name. For example, for a database named `PRODUCTIONDB` and a consumer group named `OLTP`, the `metricObjectName` would be `PRODUCTIONDB.OLTP`.

- For metrics that relate to I/O load (for example, `CG_FD_IO_LOAD` and `CG_IO_LOAD`), see the additional information relating to `CD_IO_LOAD`.

- Starting with Oracle Exadata System Software release 19.1.0, if you configured ASM-scoped security for the Oracle Automatic Storage Management (Oracle ASM) cluster used by the database, then the database name is prefixed with the Oracle ASM cluster name.

- For cumulative metrics, the metric value for a specific time period can be determined by subtracting values from different `collectionTime` periods.

- For rate metrics, the time period for the metric value is over the previous minute.

- In the metric descriptions, small I/O requests are less than or equal to 128 KB, and large I/O requests are larger than 128 KB.

- All consumer group cumulative metrics are reset to zero whenever a category, IORM, or any database resource plan is modified.

- To list the current metrics for consumer groups, use the following CellCLI command:

  ```
  CellCLI> LIST METRICCURRENT WHERE objectType = 'IORM_CONSUMER_GROUP' AND metricValue != 0 ATTRIBUTES name, metricObjectName, metricValue, collectionTime
  ```

- For Oracle ASM and all other databases, metrics are only provided for the BACKGROUND and OTHER consumer groups. The BACKGROUND consumer groups are:

  - `_ORACLE_BACKGROUND_GROUP_`: High-priority I/O requests from Oracle Database background processes
  - `_ORACLE_MEDPRIBG_GROUP_`: Medium-priority I/O requests from Oracle Database background processes
  - `_ORACLE_LOWPRIBG_GROUP_`: Low-priority I/O requests from Oracle Database background processes

### 7.3.6.3.4 Monitoring IORM with Category Metrics

Category metrics provide information about the I/O load from each category specified in the current IORM category plan.

Category metrics are identified in the `METRICCURRENT`, `METRICDEFINITION`, and `METRICHISTORY` objects as having `objectType=IORM_CATEGORY`. 7-64
The category for the metric is specified by the `metricObjectName` attribute of the `METRICCURRENT` and `METRICHISTORY` objects.

For cumulative metrics, the metric value for a specific time period can be determined by subtracting values from different `collectionTime` periods. For rate metrics, the time period for the metric value is over the previous minute.

**Example 7-10   Displaying Category Metric Definitions**

This example shows how to display the category metric definitions that are available in the Oracle Exadata System Software.

```
CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = IORM_CATEGORY
  CT_FC_IO_BY_SEC     "Number of megabytes of I/O per second for this category to flash cache"
  CT_FC_IO_RQ         "Number of IO requests issued by an IORM category to flash cache"
  CT_FC_IO_RQ_LG      "Number of large IO requests issued by an IORM category to flash cache"
  CT_FC_IO_RQ_LG_SEC  "Number of large IO requests issued by an IORM category to flash cache per second"
  CT_FC_IO_RQ_SEC     "Number of IO requests issued by an IORM category to flash cache per second"
  CT_FC_IO_RQ_SM      "Number of small IO requests issued by an IORM category to flash cache"
  CT_FC_IO_RQ_SM_SEC  "Number of small IO requests issued by an IORM category to flash cache per second"
  CT_FD_IO_BY_SEC     "Number of megabytes of I/O per second for this category to flash disks"
  CT_FD_IO_LOAD       "Average I/O load from this category for flash disks"
  CT_FD_IO_RQ_LG      "Number of large IO requests issued by an IORM category to flash disks"
  CT_FD_IO_RQ_LG_SEC  "Number of large IO requests issued by an IORM category to flash disks per second"
  CT_FD_IO_RQ_SM      "Number of small IO requests issued by an IORM category to flash disks"
  CT_FD_IO_RQ_SM_SEC  "Number of small IO requests issued by an IORM category to flash disks per second"
  CT_FD_IO_TM         "The cumulative latency of reading or writing blocks for this category from flash disks"
  CT_FD_IO_TM_LG      "The cumulative latency of reading or writing large blocks for this category from flash disks"
  CT_FD_IO_TM_LG_RQ   "The rate which is the average latency of reading or writing large blocks per request for this category from flash disks"
  CT_FD_IO_TM_RQ      "The rate which is the average latency of reading or writing blocks per request for this category from flash disks"
  CT_FD_IO_TM_SM      "The cumulative latency of reading or writing small blocks for this category from flash disks"
  CT_FD_IO_TM_SM_RQ   "The rate which is the average latency of reading or writing small blocks per request for this category from flash disks"
  CT_FD_IO_UTIL       "Percentage of flash resources used by this category"
```
CT_FD_IO_UTIL_LG        "Percentage of flash resources utilized by large requests from this category"
CT_FD_IO_UTIL_SM        "Percentage of flash resources utilized by small requests from this category"
CT_FD_IO_WT_LG          "IORM wait time for large IO requests issued to flash disks by an IORM category"
CT_FD_IO_WT_LG_RQ       "Average IORM wait time per request for large IO requests issued to flash disks by an IORM category"
CT_FD_IO_WT_SM          "IORM wait time for small IO requests issued to flash disks by an IORM category"
CT_FD_IO_WT_SM_RQ       "Average IORM wait time per request for small IO requests issued to flash disks by an IORM category"
CT_IO_BY_SEC            "Number of megabytes of I/O per second for this category to hard disks"
CT_IO_LOAD              "Average I/O load from this category for hard disks"
CT_IO_RQ_LG             "Number of large IO requests issued by an IORM category to hard disks"
CT_IO_RQ_LG_SEC         "Number of large IO requests issued by an IORM category to hard disks per second"
CT_IO_RQ_SM             "Number of small IO requests issued by an IORM category to hard disks"
CT_IO_RQ_SM_SEC         "Number of small IO requests issued by an IORM category to hard disks per second"
CT_IO_TM_LG             "The cumulative latency of reading or writing large blocks for this category from hard disks"
CT_IO_TM_LG_RQ          "The rate which is the average latency of reading or writing large blocks per request for this category from hard disks"
CT_IO_TM_SM             "The cumulative latency of reading or writing small blocks for this category from hard disks"
CT_IO_TM_SM_RQ          "The rate which is the average latency of reading or writing small blocks per request for this category from hard disks"
CT_IO_UTIL_LG           "Percentage of disk resources utilized by large requests from this category"
CT_IO_UTIL_SM           "Percentage of disk resources utilized by small requests from this category"
CT_IO_WT_LG             "IORM wait time for large IO requests issued to hard disks by an IORM category"
CT_IO_WT_LG_RQ          "Average IORM wait time per request for large IO requests issued to hard disks by an IORM category"
CT_IO_WT_SM             "IORM wait time for small IO requests issued to hard disks by an IORM category"
CT_IO_WT_SM_RQ          "Average IORM wait time per request for small IO requests issued to hard disks by an IORM category"

Note the following additional details:

• In the metric descriptions, small I/O requests are less than or equal to 128 KB, and large I/O requests are larger than 128 KB.
• The unit of measurement for the wait metrics is milliseconds. The wait metrics have metric names starting with CD_IO_WT_.

Chapter 7
Monitoring Oracle Exadata System Software Components
7-66
• For metrics that relate to I/O load (for example, CT_FD_IO_LOAD and CT_IO_LOAD), see the additional information relating to CD_IO_LOAD.

• All category cumulative metrics are reset to zero whenever a category, IORM, or any database resource plan is modified.

• To list the category metric history for an interdatabase plan, use the following CellCLI command:

  CellCLI> LIST METRICHISTORY WHERE objectType = 'IORM_CATEGORY' AND metricValue != 0 ATTRIBUTES name, metricObjectName, metricValue, collectionTime

• Category metrics are also provided for the following internally-generated and automatically-managed categories:
  – _ASM_: Oracle ASM-related I/Os
  – _ORACLE_BG_CATEGORY_: High-priority I/Os issued by Oracle Database background processes
  – _ORACLE_MEDPRIBG_CATEGORY_: Medium-priority I/Os issued by Oracle Database background processes
  – _ORACLE_LOWPRIBG_CATEGORY_: Low-priority I/Os issued by Oracle Database background processes

7.3.6.3.5 Monitoring IORM Utilization

You can use metrics to monitor IORM utilization.

When OLTP and DSS workloads share Oracle Exadata Storage Servers, IORM determines whether to optimize for low latency or high throughput. To optimize for low latency, the concurrency of large I/O requests is reduced to ensure that I/O bandwidth is not saturated. To optimize for high throughput, each Oracle Exadata Storage Server must handle many concurrent large I/O requests, allowing the storage to be fully utilized while applying optimization algorithms. However, when a cell has many concurrent large I/O requests, average latency may increase because each I/O is queued behind other I/Os.

The utilization metrics for I/O requests from a database, pluggable database (PDB), or consumer group corresponds to the amount of time that the database, PDB, or consumer group utilized the storage server. Large I/O requests utilize the storage server more than small I/O requests. The following are the utilization metrics for determining IORM optimization:

• CG_IO_UTIL_LG
• CG_IO_UTIL_SM
• PDB_IO_UTIL_LG
• PDB_IO_UTIL_SM
• CT_IO_UTIL_LG
• CT_IO_UTIL_SM
• DB_IO_UTIL_LG
• DB_IO_UTIL_SM
By comparing the I/O resource consumption with the I/O resource allocations, the database administrator can determine if IORM should be tuned for latency or throughput, or if a balanced approach is optimal. The IORM metric, IORM_MODE, shows the mode for IORM. The metric value ranges between 1 and 3. The following are the definitions for the values:

- **1** means the cell IORM objective is set to low_latency.
- **2** means the cell IORM objective is set to balanced.
- **3** means the cell IORM objective is set to high_throughput.

A value in between 1 and 2, or between 2 and 3, indicates that the IORM objective changed in the metric period, and the precise value indicates proximity to a given objective. It is also indicative of a constantly-changing mix of workloads.

**7.3.6.4 What to Look For When Monitoring I/O Resource Management (IORM)**

**I/O Latency**

Issues with IORM typically result in increased I/O latency. This is usually characterized by higher latency in the cell single block physical read database wait event, and in some cases the cell smart table scan database wait event. If these database wait events are significant, and there is no corresponding latencies in the storage servers that are associated with the flash or disk devices, then this may indicate that IORM is throttling the workload.

To confirm that IORM throttling is occurring:

- In the Top Databases by Requests - Details section of the AWR report, review the Queue Time columns, which show the average amount of time that IORM spent throttling IO requests for the database.
- Review the cell metrics for IORM wait times: DB_IO WT_SM_RQ, DB_IO WT_LG_RQ, PDB_IO WT_SM_RQ, PDB_IO WT_LG_RQ and CG_IO WT_SM_RQ.

You can use IORM database statistics and the AWR report to understand the I/O workload as a whole. You can use Exadata metrics to further understand I/O consumption by each category, database, or consumer group. By analyzing statistics and metrics, you can understand which category, database, pluggable database (PDB), or consumer group is not using its resource allocation and which is exceeding its resource allocation.

If the wait times are small or zero, then the plan allocation is sufficient. If the wait times are large, then the plan allocation is insufficient. If the wait times due to high I/O latency result in unacceptable performance, then the IORM plan can be adjusted to give a larger allocation, or additional storage servers may be required to deliver the required I/O resources.
Using IOSTAT

Device utilization and I/O service time statistics monitored using `iostat` (which is collected by ExaWatcher) are unreliable.

The following is stated in the Linux man page for `iostat`:

- `svctm` - The average service time (in milliseconds) for I/O requests that were issued to the device. Warning! Do not trust this field any more. This field will be removed in a future `sysstat` version.

Since the utilization computation depends upon the I/O service time, it is also unreliable.

To monitor the actual I/O utilization for a cell disk, database, pluggable database or consumer group, use the corresponding IORM metrics.

7.3.7 Monitoring Cell Disk I/O

When database performance issues are related to I/O load on the Exadata storage servers, typically there will be increased latencies in the I/O-related wait events, and increased database time in the User I/O or System I/O wait classes. If the increased database latencies are due to the performance of the Exadata storage servers, then the increased latencies will also be visible in the cell-side statistics. If you have comprehensive baseline statistics for periods when the system is performing well, then you can compare the baseline statistics with other observations to identify differences and explain the situation.

For example, if Oracle Database reports increased cell single block physical read latencies, you can then check the statistics from the storage servers and compare them to a baseline to determine if the cause is increased latency on flash devices, more disk I/O, or some other cause. If the statistics show an increase in disk I/O requests, that may be related to a change in Exadata Smart Flash Cache, which would prompt you to review the Exadata Smart Flash Cache statistics. On the other hand, if the statistics show a latency increase for small reads from flash, then this may be caused by a change in the I/O pattern, and understanding the type of I/O that has increased (small reads, small writes, large reads, or large writes) can help to drive further investigation.

Typically, in an Exadata environment, an even workload distribution is expected across all cells and all disks. If one cell or one disk is doing more work than the others, then you should investigate further, as that cell or disk has the potential of slowing down the entire system.

- Monitoring Cell Disk I/O Using AWR
- Monitoring Cell Disk I/O Using Database Statistics and Wait Events
- Monitoring Cell Disk I/O Using Exadata Metrics
- What to Look For When Monitoring Cell Disk I/O

7.3.7.1 Monitoring Cell Disk I/O Using AWR

The following sections in the Automatic Workload Repository (AWR) report are particularly useful for understanding I/O load on Exadata:

- Disk Activity
• Exadata Resource Statistics
• Exadata IO Reasons
• Internal IO Reasons

Often, to better understand characteristics about the I/O load, the statistics from these sections can be correlated with other sections in the AWR report.

Disk Activity

The Disk Activity section provides a high-level summary for potential sources of disk activity. The Disk Activity section is located in the AWR report under Exadata Statistics > Performance Summary.

Figure 7-20  AWR Report: Disk Activity

<table>
<thead>
<tr>
<th>I/O per second</th>
<th>% Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redo log writes</td>
<td>26,766.87</td>
</tr>
<tr>
<td>Smart Scans (estd)</td>
<td></td>
</tr>
<tr>
<td>Flash Cache misses (OLTP)</td>
<td>941.25</td>
</tr>
<tr>
<td>Flash Cache read skips</td>
<td></td>
</tr>
<tr>
<td>Flash Cache write skips</td>
<td>41.98</td>
</tr>
<tr>
<td>Disk writer writes</td>
<td>0.90</td>
</tr>
<tr>
<td>Scrub IO</td>
<td>1.14</td>
</tr>
<tr>
<td>H/12.5T</td>
<td>16.75</td>
</tr>
</tbody>
</table>

High I/O load or a substantial change in the pattern of disk activity may prompt further investigation. Possible causes include:

- Redo log write — result in disk writes when redo is written to disk. When using Exadata Smart Flash Log, note that redo is written to both Exadata Smart Flash Log and the online redo log file. Also, Oracle Exadata System Software release 20.1 adds a further optimization, known as Smart Flash Log Write-Back, that uses Exadata Smart Flash Cache in Write-Back mode instead of disk storage. For further details, review Database Redo Activity and Smart Flash Log in the AWR report.

- Smart Scans — result in disk reads for requests that are not satisfied using Exadata Smart Flash Cache. These are typically large reads. For further details, review Smart IO in the AWR report.

- Flash Cache misses — result in disk reads when requested data does not exist in Exadata Smart Flash Cache. These are typically small reads. For further details, review Flash Cache Misses in the AWR report.

- Flash Cache read skips — result in disk reads when requested data is not eligible for Exadata Smart Flash Cache. For further details, review Flash Cache User Reads - Skips in the AWR report.
• Flash Cache write skips or Flash Cache LW rejections — results in disk writes when data is not eligible for Exadata Smart Flash Cache. For further details, review Flash Cache User Writes - Skips and Flash Cache User Writes - Large Write Rejects in the AWR report.

• Disk writer writes — result in disk writes when data from Exadata Smart Flash Cache in Write-Back mode is persisted to disk. For further details, review Flash Cache Internal Writes in the AWR report.

• Scrub IO — occurs when Oracle Exadata System Software automatically inspects and repairs the hard disks. Scrub I/O is performed periodically when the hard disks are idle, and mostly results in large disk reads, which should be throttled automatically if the disk becomes I/O bound.

The specific causes listed in this section are subject to the version of Oracle Database being used.

Exadata Resource Statistics

The Exadata Resource Statistics section contains many statistics and is organized into several sub-sections. Primarily, the statistics enumerate the I/O occurring on the storage servers using information from the storage server operating system (OS) and Oracle Exadata System Software. From the OS, it includes statistics relating to I/Os per second (IOPS), throughput, utilization, service time and wait time. These statistics are equivalent to the statistics shown by the `iostat` command. From the Oracle Exadata System Software, it includes statistics relating to IOPS, throughput, and latency, which are also broken down by small reads, small writes, large reads, and large writes. These statistics are based on the cell disk metrics.

The statistics are aggregated by device type, and then by cell or disk. The device type aggregation ensures comparison across the same device type, as different device types are expected to have different performance characteristics. The statistics are presented two ways. Firstly, they are presented to enable outlier analysis. They are also organized to show the ‘Top N’ cells or disks for a specific statistic.

The outlier analysis presentation allows you to quickly see the statistics aggregated across all storage servers, by cell, and by disk. The display also includes the statistical mean, standard deviation, and normal range. The normal range is based on the mean and standard deviation, not the observed low and high values. For cells, the normal range is the range of values that are plus or minus one standard deviation from the mean. For disks, the normal range is the range of values that are plus or minus three standard deviations from the mean. If there are cells or disks that fall outside the normal range, then they are reported as outliers. This simple outlier analysis is designed to highlight areas for further investigation. However, based on the number of cells or disks in use, and the value of the standard deviation, the outlier analysis may not identify outliers in all cases.

The ‘Top N’ presentation simply shows the top ranked cells or disks for a specific statistic. This presentation enables you to identify cells or disks that perform more or less work than the others. By using this presentation, you can potentially identify outliers that are not identified by the outlier analysis. Also highlighted in these sections are cells or disks that exceed the expected maximum IOPS for the device or the expected maximum throughput for the device.

The following list outlines the sub-sections in the Exadata Resource Statistics section of the AWR report:
Exadata Outlier Summary — displays a summary of outliers identified in the various outlier sub-sections.

Exadata OS Statistics Outliers — contains outlier sub-sections for cells and disks based on OS statistics including, IOPS, throughput (MB/s), utilization percentage, service time, wait time, and CPU utilization per cell.

Exadata Cell Server Statistics — contains outlier sub-sections for cells and disks based on cell disk metrics, including IOPS, throughput (MB/s), and latency. The statistics are further categorized by I/O type; that is, small read, small write, large read, or large write.

Exadata Outlier Details — displays detailed information for the identified outliers, along with other statistics related to the outlier.

Exadata OS Statistics Top — contains 'Top N' sub-sections for cells and disks based on OS statistics, including IOPS, latency, and CPU utilization.

Exadata Cell Server Statistics Top — contains 'Top N' sub-sections for cells and disks based on cell disk metrics, including IOPS, throughput (MB/s), and latency. The statistics are further categorized by I/O type; that is, small read, small write, large read, or large write.

The following example shows two of the Exadata Cell Server Statistics outlier sub-sections in the AWR report. The example highlights that hard disk throughput (IOPS) exceeds the expected maximum, and that a specific disk that is performing substantially more small reads than other disks.

Figure 7-21  AWR Report: Exadata Cell Server IOPS Statistics Outliers

Exadata IO Reasons

When the database sends an I/O request to Exadata, the request is tagged with information that includes the reason for the I/O. This information is aggregated in the Exadata IO Reasons sections of the AWR report and allows you to understand the reasons for performing I/O.

The AWR report contains sub-sections that display the Top IO Reasons by Requests and the Top IO Reasons by MB (throughput). Later versions of the AWR report, further
break down the Top IO Reasons by Requests to categorize read requests from flash, write requests to flash, read requests from disk, and write requests to disk.

The following example shows Top IO Reasons by Requests. The example output is typical of a well-performing system, with a high proportion of I/O that is associated with Smart Scan, and similar I/O profiles across all of the storage servers.

Figure 7-22  AWR Report: Top IO Reasons by Request

Top IO Reasons by Requests

- The top IO reasons by requests per cell are displayed
- Only reasons with over 1% of IO requests for each cell are displayed
- At most 10 reasons are displayed per cell
- %Cell - the percentage of IO requests on the cell due to the IO reason
- Ordered by Cell Name, Requests Value desc

<table>
<thead>
<tr>
<th>Cell Name</th>
<th>IO Reason</th>
<th>%Cell</th>
<th>Total</th>
<th>per Sec</th>
<th>Total MB</th>
<th>per Sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (3)</td>
<td>smart scan</td>
<td>67.75</td>
<td>606,238,313</td>
<td>169,387.63</td>
<td>38,272,870</td>
<td>10,693.73</td>
</tr>
<tr>
<td></td>
<td>buffer cache reads</td>
<td>25.77</td>
<td>230,585,955</td>
<td>64,427.48</td>
<td>3,035,528</td>
<td>848.15</td>
</tr>
<tr>
<td></td>
<td>aged writes by dbwr</td>
<td>4.54</td>
<td>40,626,048</td>
<td>11,351.23</td>
<td>351,331</td>
<td>98.16</td>
</tr>
<tr>
<td>dbm0celadm01</td>
<td>smart scan</td>
<td>67.65</td>
<td>199,997,190</td>
<td>55,880.75</td>
<td>12,614,069</td>
<td>3,524.47</td>
</tr>
<tr>
<td></td>
<td>buffer cache reads</td>
<td>25.74</td>
<td>76,091,088</td>
<td>21,260.43</td>
<td>999,835</td>
<td>279.36</td>
</tr>
<tr>
<td></td>
<td>aged writes by dbwr</td>
<td>4.54</td>
<td>13,414,343</td>
<td>3,748.07</td>
<td>116,088</td>
<td>32.44</td>
</tr>
<tr>
<td>dbm0celadm02</td>
<td>smart scan</td>
<td>68.02</td>
<td>203,512,861</td>
<td>56,863.05</td>
<td>12,838,453</td>
<td>3,587.16</td>
</tr>
<tr>
<td></td>
<td>buffer cache reads</td>
<td>25.74</td>
<td>77,029,155</td>
<td>21,522.54</td>
<td>1,015,686</td>
<td>283.79</td>
</tr>
<tr>
<td></td>
<td>aged writes by dbwr</td>
<td>4.53</td>
<td>13,541,707</td>
<td>3,783.66</td>
<td>116,890</td>
<td>32.66</td>
</tr>
<tr>
<td>dbm0celadm03</td>
<td>smart scan</td>
<td>67.58</td>
<td>202,728,262</td>
<td>56,643.83</td>
<td>12,820,347</td>
<td>3,582.10</td>
</tr>
<tr>
<td></td>
<td>buffer cache reads</td>
<td>25.82</td>
<td>77,465,712</td>
<td>21,644.51</td>
<td>1,020,066</td>
<td>285.00</td>
</tr>
<tr>
<td></td>
<td>aged writes by dbwr</td>
<td>4.56</td>
<td>13,669,998</td>
<td>3,819.50</td>
<td>118,352</td>
<td>33.07</td>
</tr>
</tbody>
</table>

Internal IO Reasons

If Internal IO is among the Top IO Reasons reported, then the AWR report will include a section that summarizes the Internal IO Reasons:

Figure 7-23  AWR Report: Internal IO Reasons

Internal IO Reasons

- The following are possible reasons for Internal IO
- The values displayed are the total IOs over all cells
- Population Writes for columnar cache - top level call that can result in multiple writes to flash

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Total</th>
<th>per Sec</th>
<th>Reads/s</th>
<th>Writes/s</th>
<th>Total</th>
<th>per Sec</th>
<th>Reads/s</th>
<th>Writes/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal IO</td>
<td>35,768,322</td>
<td>18,230.54</td>
<td>3,808.53</td>
<td>14,422.01</td>
<td>3,796.20</td>
<td>14,024.99</td>
<td>12.33</td>
<td>397.02</td>
</tr>
<tr>
<td>Disk Writer reads</td>
<td>7,446,408</td>
<td>3,795.31</td>
<td>3,795.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk Writer writes</td>
<td>774,794</td>
<td>394.90</td>
<td>394.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metadata</td>
<td>27,430,769</td>
<td>13,981.02</td>
<td>13,981.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Possible causes of internal I/O include:

- Disk Writer reads — results in flash reads when the disk writer reads from Exadata Smart Flash Cache in Write-Back mode to persist data to disk. For further details, review Flash Cache Internal Reads in the AWR report.

- Disk Writer writes — results in disk writes when the disk writer persists data to disk from Exadata Smart Flash Cache in Write-Back mode. For further details, review Flash Cache Internal Reads in the AWR report.

- Population — results in flash writes when requested data is read into Exadata Smart Flash Cache. When the data is read from disk, it also populates Exadata Smart Flash Cache. This is often correlated with flash cache misses. For further details, review Flash Cache User Reads and Flash Cache Internal Writes in the AWR report.

- Metadata — results in flash writes when new data is written to Exadata Smart Flash Cache in Write-Back mode. This is often due to first writes. For further details, review Flash Cache User Writes in the AWR report.

The specific causes listed in this section are subject to the version of Oracle Database being used.

### 7.3.7.2 Monitoring Cell Disk I/O Using Database Statistics and Wait Events

The following table describes various database statistics that are useful for monitoring cell disk I/O. The statistics are available in various dynamic performance views, including V$SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>physical read IO requests</td>
<td>Number of physical read requests issued by the database.</td>
</tr>
<tr>
<td>physical read requests optimized</td>
<td>Total number of read requests satisfied by using Exadata Smart Flash Cache, and read requests avoided by using storage index or columnar cache.</td>
</tr>
<tr>
<td>physical read total bytes</td>
<td>Total amount of I/O bytes for reads issued by the database, whether or not the read was offloaded to the storage servers.</td>
</tr>
<tr>
<td>physical read total bytes optimized</td>
<td>Total number of bytes read from Exadata Smart Flash Cache, and bytes avoided by using storage index or columnar cache.</td>
</tr>
<tr>
<td>physical read total IO requests</td>
<td>Total number of read requests issued by the database for all activity including application, backup, recovery, and other utilities.</td>
</tr>
<tr>
<td>physical write IO requests</td>
<td>Number of physical write requests issued by the database.</td>
</tr>
<tr>
<td>physical write total bytes</td>
<td>Total number of I/O bytes for writes issued by the database for all activity.</td>
</tr>
<tr>
<td>physical write total bytes optimized</td>
<td>Total number of bytes written to Exadata Smart Flash Cache. These bytes are synchronized to disk in a lazy manner.</td>
</tr>
<tr>
<td>physical write total IO requests</td>
<td>Total number of write requests issued by the database for all activity, including application, backup, recovery and other utilities.</td>
</tr>
</tbody>
</table>
Additional database I/O statistics are also contained in the following V$STAT views, which are also used in the IOStat sections of the AWR report.

- **V$STAT_FUNCTION** — displays disk I/O statistics for database functions, such as the LGWR and DBWR.
- **V$STAT_FUNCTION_DETAIL** — displays disk I/O statistics for database functions, broken down by file type.

The following table describes database wait events that are useful for monitoring cell disk I/O. The wait events are visible in various dynamic performance views, including V$SESSION, V$SYSTEM_EVENT and V$SESSION_EVENT, and may be displayed in the Wait Event sections of the AWR report.

<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell list of blocks physical read</td>
<td>This wait event occurs during recovery or during buffer pre-fetching (rather than performing multiple single-block reads). It is used to monitor database blocks that must be changed as part of recovery and are read in parallel for the database. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the number of blocks processed during the I/O read operation. This wait event is equivalent to db file parallel read for a cell.</td>
</tr>
<tr>
<td>cell multiblock physical read</td>
<td>The wait time is the actual time it takes to do all of the I/Os when reading multiple data blocks. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the total number of bytes processed during the I/O read operation. This wait event is equivalent to db file scattered read for a cell.</td>
</tr>
<tr>
<td>cell single block physical read</td>
<td>The wait time is the actual time it takes to do the I/O while a sequential read for the database is performed. The P1, P2, and P3 values associated with this event in V$SESSION identify the cell hash number, disk hash number, and the number of bytes processed during the I/O read operation. This wait event is equivalent to db file sequential read for a cell.</td>
</tr>
</tbody>
</table>
Wait Event | Description
---|---
cell interconnect retransmit during physical read | This wait event appears during retransmission for an I/O of a single-block or multiblock read. The cell hash number in the P1 column in the V$SESSION_WAIT view is the same cell identified for cell single block physical read and cell multiblock physical read. The P2 column contains the subnet number to the cell, and the P3 column contains the number of bytes processed during the I/O read operation.

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.

### 7.3.7.3 Monitoring Cell Disk I/O Using Exadata Metrics

Cell disk metrics provide information about the I/O load for cell disks, such as the number of large blocks read from a cell disk.

Cell disk metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having `objectType=CELLDISK`.

The cell disk for the metric is specified by the `metricObjectName` attribute of the METRICCURRENT and METRICHISTORY objects.

For cumulative metrics, the metric value for a specific time period is determined by subtracting values from different `collectionTime` periods. For rate metrics, the time period for the metric value is over the previous minute.

#### Example 7-11  Displaying Cell Disk Metric Definitions

This example shows how to display the cell disk metric definitions that are available in the Oracle Exadata System Software.

```
CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = CELLDISK
CD_BY_FC_DIRTY          "Number of unflushed megabytes cached in FLASHCACHE on a cell disk"
CD_IO_BY_R_LG           "Number of megabytes read in large blocks from a cell disk"
CD_IO_BY_R_LG_SEC       "Number of megabytes read in large blocks per second from a cell disk"
CD_IO_BY_R_SCRUB        "Number of megabytes read from a cell disk by the scrubbing job"
CD_IO_BY_R_SCRUB_SEC    "Number of megabytes read per second from a cell disk by the scrubbing job"
CD_IO_BY_R_SM           "Number of megabytes read in small blocks from a cell disk"
CD_IO_BY_R_SM_SEC       "Number of megabytes read in small blocks per second from a cell disk"
CD_IO_BY_W_LG           "Number of megabytes written in large blocks to a cell disk"
CD_IO_BY_W_LG_SEC       "Number of megabytes written in large blocks per second to a cell disk"
```
CD_IO_BY_W_SM          "Number of megabytes written in small blocks to a cell disk"
CD_IO_BY_W_SM_SEC      "Number of megabytes written in small blocks per second to a cell disk"
CD_IO_ERRS            "Number of IO errors on a cell disk"
CD_IO_ERRS_MIN        "Number of IO errors on a cell disk per minute"
CD_IO_ERRS_SCRUB      "Number of IO errors hit by the scrubbing job on a cell disk"
CD_IO_LOAD            "Average I/O load for the cell disk"
CD_IO_RQ_R_LG         "Number of requests to read large blocks from a cell disk"
CD_IO_RQ_R_LG_SEC     "Number of requests to read large blocks per second from a cell disk"
CD_IO_RQ_R_SCRUB      "Number of requests to read from a cell disk by the scrubbing job"
CD_IO_RQ_R_LG_RQ      "Number of requests to read per second from a cell disk by the scrubbing job"
CD_IO_RQ_R_SM         "Number of requests to read small blocks from a cell disk"
CD_IO_RQ_R_SM_SEC     "Number of requests to read small blocks per second from a cell disk"
CD_IO_RQ_W_LG         "Number of requests to write large blocks to a cell disk"
CD_IO_RQ_W_LG_SEC     "Number of requests to write large blocks per second to a cell disk"
CD_IO_RQ_W_SM         "Number of requests to write small blocks to a cell disk"
CD_IO_RQ_W_SM_SEC     "Number of requests to write small blocks per second to a cell disk"
CD_IO_ST_RQ           "Average service time per request for small IO requests to a cell disk"
CD_IO_TM_R_LG         "Cumulative latency of reading large blocks from a cell disk"
CD_IO_TM_R_LG_RQ      "Average latency of reading large blocks per request from a cell disk"
CD_IO_TM_R_SM         "Cumulative latency of reading small blocks from a cell disk"
CD_IO_TM_R_SM_RQ      "Average latency of reading small blocks per request from a cell disk"
CD_IO_TM_W_LG         "Cumulative latency of writing large blocks to a cell disk"
CD_IO_TM_W_LG_RQ      "Average latency of writing large blocks per request to a cell disk"
CD_IO_TM_W_SM         "Cumulative latency of writing small blocks to a cell disk"
CD_IO_TM_W_SM_RQ      "Average latency of writing small blocks per request to a cell disk"
CD_IO_UTIL            "Percentage of disk resources utilized for the cell disk"
CD_IO_UTIL_LG         "Percentage of disk resources utilized by large requests for the cell disk"
CD_IO_UTIL_SM         "Percentage of disk resources utilized by small requests for the cell disk"
Note the following additional details:

- **CD_IO_LOAD** represents the average I/O load for the cell disk. 

  I/O load specifies the length of the disk queue. It is similar to iostat's `avgqu-sz`, but I/O load is a weighted value depending on the type of disk:
  
  - For hard disks, a large I/O has three times the weight of a small I/O.
  - For flash disks, large and small I/Os have the same weight.

  Corresponding metrics are also available for each database (**DB_IO_LOAD**), pluggable database (PDB) (**PDB_IO_LOAD**), IORM category (**CT_IO_LOAD**), and consumer group (**CG_IO_LOAD**).

- **CD_IO_ST_RQ** cannot be used for flash devices, as explained in Using IOSTAT.

- **CD_IO_UTIL** is similar to `%util` of iostat for hard disks, but not for flash devices as mentioned in Using IOSTAT. For flash devices, it is a percentage of the maximum I/O bandwidth for your system, as specified in the product data sheet.

  Because this metric is computed by IORM, it is also available per database, PDB, and consumer group.

- The unit of measurement for the latency metrics is microseconds. The latency metrics have metric names starting with **CD_IO_TM_**.

### 7.3.7.4 What to Look For When Monitoring Cell Disk I/O

#### Imbalances

In an Exadata environment, an even load distribution is expected across all cells or disks. However, there are situations that may cause an imbalance.

A load imbalance may be due to characteristics of the workload, such as:

- Repeated small-table scans — This is often caused by a table on the right side of a nested loop join. Since the small table may only reside on a few disks or cells, repeated access means reading from a small set of devices, which may be flash devices if the data resides in Exadata Smart Flash Cache. To address the imbalance, you can identify affected SQL statements and review their execution plans. You can also consider the physical organization of the affected segments.

- Repeated control file reads — Control file reads may be caused by queries against some database dynamic performance views. The control file is small and may only reside on a few disks or cells, so repeated access means reading from a small set of devices, usually flash. Repeated control file reads are evident in the IOSTat by File Type section of the AWR report and using the control file sequential read wait event.

  To understand the cause of repeated control file reads, you can use the Exadata IO Reasons section of the AWR report to identify cells servicing many control file reads and correlate this with statistics from the Top Databases section to identify databases issuing the control file reads. You can also review Active Session History (ASH) to identify SQL statements that are experiencing waits for control file sequential read.

  Beware that control file reads may account for a disproportionate amount of small I/O on a lightly-loaded system. In this case, you can safely ignore the imbalance.
• Repeated reads of a small segment, such as a LOB segment — To address
this kind of imbalance, review database statistics for segment, or examine ASH
to identify the responsible SQL commands. Then, examine the SQL and the
surrounding application logic to see if the repeated reads can be avoided.

• Repeated ASM metadata access — This is often caused by database queries
related to space usage, which require access to ASM metadata. The ASM
metadata is small and may only reside on a few disks or cells, so repeated access
may show up as an imbalance. This can show up in the Exadata IO Reasons
sections of the AWR report with reasons that are prefixed with ASM, such as
ASM CACHE IO. You can use the Exadata IO Reasons section to identify affected
cells and correlate this with statistics from the Top Databases sections to confirm
ASM as the source of the I/O. To address the imbalance, review the need for and
frequency of the ASM space usage queries that access the metadata.

Apart from workload-related causes, imbalances may also be caused by:

• Uneven data distribution — Hot spots may occur when some disks contain more
or less data than others. To check for balanced data distribution, you can query the
V$ASM_DISK_IOSTAT view before and after running a SQL query that contains large
full-table scans. In this case, the statistics in the read column and the read_bytes
column should be approximately the same for all disks in the disk group. You can
also check for balanced data distribution by using the script available in My Oracle
Support document 367445.1.

• Asymmetric grid disk configurations — If the grid disks are sized differently
on different cells, then different amounts of data may reside on each cell,
resulting in imbalance due to uneven data distribution. For this reason, symmetric
configuration is recommended across all storage servers in a system.

• Failures or recovery from failures — Some processing cannot spread evenly
across the remaining disks or cells. For example, if a flash card fails, the cell
containing the card has less available flash cache space, and this may show up as
an imbalance when the flash cache statistics are compared with the other cells.

An imbalance often shows up in different statistics across the AWR report. By
correlating different sections in the AWR report you may gain a deeper understanding
of the imbalance.

High I/O Load

When database performance issues are related to I/O load on the Exadata storage
servers, typically there will be increased latencies in the I/O related wait events, and
increased database time in the user I/O or system I/O wait classes. When dealing
with high I/O load, first understand the composition of the I/O load. Based on the
composition of the I/O load you my be directed to different statistics to gain a deeper
understanding.

In the AWR report, a good starting point is the Disk Activity section, which provides
a high-level summary for potential sources of disk activity. The Disk Activity section is
located in the AWR report under Exadata Statistics > Performance Summary. See Disk
Activity.

Internal I/O

In the AWR report, if internal I/O is reported among the Top IO Reasons, then the
AWR report also includes a section that summarizes the Internal IO Reasons. Use
this section to understand the composition of internal I/O. Based on the composition
of the internal I/O load you may be directed to different statistics to gain a deeper understanding. See Internal IO Reasons.

7.3.8 Monitoring Grid Disks

Grid disk metrics provide information about the I/O load for grid disks, such as the number of large blocks read from a grid disk.

Grid disk metrics are identified in the METRICCURRENT, METRICDEFINITION, and METRICHISTORY objects as having objectType=GRIDDISK.

The grid disk for the metric is specified by the metricObjectName attribute of the METRICCURRENT and METRICHISTORY objects.

For cumulative metrics, the metric value for a specific time period is determined by subtracting values from different collectionTime periods. For rate metrics, the time period for the metric value is over the previous minute.

Example 7-12   Displaying Grid Disk Metric Definitions

This example shows how to display the grid disk metric definitions that are available in the Oracle Exadata System Software.

CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = GRIDDISK
  GD_IO_BY_R_LG                  "Number of megabytes read in large blocks from a grid disk"
  GD_IO_BY_R_LG_SEC             "Number of megabytes read in large blocks per second from a grid disk"
  GD_IO_BY_R_SCRUB              "Number of megabytes read per second from a grid disk by the scrubbing job"
  GD_IO_BY_R_SCRUB_SEC         "Number of megabytes read per second from a grid disk by the scrubbing job"
  GD_IO_BY_R_SM                 "Number of megabytes read in small blocks from a grid disk"
  GD_IO_BY_R_SM_SEC            "Number of megabytes read in small blocks per second from a grid disk"
  GD_IO_BY_W_LG                 "Number of megabytes written in large blocks to a grid disk"
  GD_IO_BY_W_LG_SEC            "Number of megabytes written in large blocks per second to a grid disk"
  GD_IO_BY_W_SM                 "Number of megabytes written in small blocks to a grid disk"
  GD_IO_BY_W_SM_SEC            "Number of megabytes written in small blocks per second to a grid disk"
  GD_IO_ERRS                   "Number of IO errors on a grid disk"
  GD_IO_ERRS_MIN               "Number of IO errors on a grid disk per minute"
  GD_IO_RQ_R_LG                "Number of requests to read large blocks from a grid disk"
  GD_IO_RQ_R_LG_SEC           "Number of requests to read large blocks per second from a grid disk"
large blocks per second from a grid disk"  
GD_IO_RQ_R_SCRUB  
"Number of requests to read  
from a grid disk by the scrubbing job"  
GD_IO_RQ_R_SCRUB_SEC  
"Number of requests to read  
per second from a grid disk by the scrubbing job"  
GD_IO_RQ_R_SM  
"Number of requests to read  
small blocks from a grid disk"  
GD_IO_RQ_R_SM_SEC  
"Number of requests to read  
small blocks per second from a grid disk"  
GD_IO_RQ_W_LG  
"Number of requests to write  
large blocks to a grid disk"  
GD_IO_RQ_W_LG_SEC  
"Number of requests to write  
large blocks per second to a grid disk"  
GD_IO_RQ_W_SM  
"Number of requests to write  
small blocks to a grid disk"  
GD_IO_RQ_W_SM_SEC  
"Number of requests to write  
small blocks per second to a grid disk"  
GD_NVM_READ_RETRIES  
"Number of read retries for a  
non-volatile memory (NVM) grid disk"  
GD_SP_BY_ALLOCATED  
"Allocated physical space for  
grid disk in bytes"  
GD_SP_IO_BY_PARTIAL  
"Bytes returned by partial I/Os"  
GD_SP_IO_BY_PARTIAL_SEC  
"Bytes returned by partial I/Os  
per second"  
GD_SP_IO_BY_REDIRECTED  
"Sparse bytes redirected to  
original data block"  
GD_SP_IO_BY_REDIRECTED_SEC  
"Sparse bytes redirected to  
orignal data block per second"  
GD_SP_IO_QR_PARTIAL  
"Number of IO requests that  
rerturned partial data"  
GD_SP_IO_QR_PARTIAL_SEC  
"Number of IO requests that  
rerturned partial data per second"  
GD_SP_PRCT_ALLOCATED  
"Allocated physical space for  
grid disk by percentage"

Note the following additional details:

- GD_SP_PRCT_ALLOCATED tracks the allocated physical space for the grid disk, and is  
presented as a percentage. By default, a warning alert is generated when space  
usage is more than 95%, and a critical alert is generated when space usage  
reaches 99%. To lower the space usage, you can increase the size of the grid  
disk, or drop some of the data on the grid disk.

7.3.9 Monitoring Host Interconnect Metrics

Host interconnect metrics provide information about the I/O transmission for hosts that  
access cell storage.

Host interconnect metrics are identified in the METRICCURRENT, METRICDEFINITION, and  
METRICHISTORY objects as having objectType=HOST_INTERCONNECT.

The host interconnect for the metric is specified by the metricObjectName attribute  
of the METRICCURRENT and METRICHISTORY objects. For cumulative metrics, the metric  
value for a specific time period can be determined by subtracting values from different
collectionTime periods. For rate metrics, the time period for the metric value is over the previous minute.

**Example 7-13  Displaying Host Interconnect Metric Definitions**

This example shows how to display the host interconnect metric definitions that are available in the Oracle Exadata System Software.

```
CellCLI> LIST METRICDEFINITION ATTRIBUTES NAME,DESCRIPTION WHERE OBJECTTYPE = HOST_INTERCONNECT
N_MB_DROP               "Number of megabytes dropped during transmission to a particular host"
N_MB_DROP_SEC           "Number of megabytes dropped during transmission per second to a particular host"
N_MB_RDMA_DROP           "Number of megabytes dropped during RDMA transmission to a particular host"
N_MB_RDMA_DROP_SEC      "Number of megabytes dropped during RDMA transmission per second to a particular host"
N_MB_RECEIVED           "Number of megabytes received from a particular host"
N_MB_RECEIVED_SEC       "Number of megabytes per second received from a particular host"
N_MB_RESENT             "Number of megabytes resent to a particular host"
N_MB_RESENT_SEC         "Number of megabytes resent per second to a particular host"
N_MB_SENT               "Number of megabytes transmitted to a particular host"
N_MB_SENT_SEC           "Number of megabytes transmitted per second to a particular host"
N_RDMA_RETRY_TM         "Latency of the retry actions during RDMA transmission to a particular host"
```

7.3.10 Monitoring RAM Cache

A random access memory (RAM) cache feature was added in Oracle Exadata System Software release 18.1, prior to the introduction of PMEM cache. The RAM cache boosts OLTP read performance at the cost of using additional storage server memory. The RAM cache acts as an extension of the database buffer cache, but it has a much smaller capacity than Exadata Smart Flash Cache or PMEM cache.

The RAM cache is not configured by default (`ramCacheMode=auto`). To use the RAM cache, it must be explicitly configured on each storage server.

- Monitoring RAM Cache Using AWR
- Monitoring RAM Cache Using Database Statistics and Wait Events

7.3.10.1 Monitoring RAM Cache Using AWR

If the RAM cache is used, there are sections in the AWR report that contain RAM cache statistics, which are similar to sections for Exadata Smart Flash Cache and PMEM cache. These include:

- Memory Cache Configuration
• Memory Cache Space Usage
• Memory Cache User Reads — Contains statistics on reads from the RAM cache that are done on behalf of the databases.
• Memory Cache Internal Writes — Contains statistics on writes to the RAM cache that are performed by Oracle Exadata System Software.

7.3.10.2 Monitoring RAM Cache Using Database Statistics and Wait Events

The following table describes database statistics that are useful for monitoring RAM cache. The statistics are available in various dynamic performance views, including V$SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell ram cache read hits</td>
<td>The number of read requests that hit the RAM cache on the cell.</td>
</tr>
</tbody>
</table>

The database wait events that are useful for monitoring the RAM cache are the same wait events that are associated with Exadata Smart Flash Cache and PMEM cache.

The use of RAM cache is typically indicated by very low latency for the cell single block physical read event.

The availability of a specific statistic or wait event is subject to the version of Oracle Database being used.

7.4 Using Exadata Monitoring Objects

• Displaying Metrics
  You can use the CellCLI LIST command to display and monitor metrics for cell objects.

• Monitoring Alerts
  You can monitor and receive notifications for alerts.

• Displaying Active Requests
  Use the LIST ACTIVEREQUEST command to display the active requests for the storage server.

7.4.1 Displaying Metrics

You can use the CellCLI LIST command to display and monitor metrics for cell objects.

You can display metric definitions, current metrics, and metric history using the LIST command.

• Displaying Metric Definitions
  Use the LIST METRICDEFINITION command to display the metric definitions for a storage server.

• Displaying Current Metrics
  Use the LIST METRICCURRENT command to display the current metric values for a storage server.
7.4.1.1 Displaying Metric Definitions

Use the LIST METRICDEFINITION command to display the metric definitions for a storage server.

A metric definition listing shows the configuration of a metric.

Example 7-14  Displaying Metric Definitions

This example shows how to display attributes for the METRICDEFINITION object.

CellCLI> LIST METRICDEFINITION CL_CPUT DETAIL

CellCLI> LIST METRICDEFINITION WHERE objectType = 'GRIDDISK'

CellCLI> LIST METRICDEFINITION WHERE name LIKE 'CD_IO_RQ.*' -
    ATTRIBUTES name, metricType, description

7.4.1.2 Displaying Current Metrics

Use the LIST METRICCURRENT command to display the current metric values for a storage server.

A current metric listing shows a set of observations on the current value of an individual metric.

Example 7-15  Displaying Current Metric Values

This example shows how to display attributes for the METRICCURRENT object.

CellCLI> LIST METRICCURRENT CL_TEMP DETAIL

<table>
<thead>
<tr>
<th>name:</th>
<th>CL_TEMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>alertState:</td>
<td>normal</td>
</tr>
<tr>
<td>collectionTime:</td>
<td>2009-12-17T15:32:25-08:00</td>
</tr>
<tr>
<td>metricObjectName:</td>
<td>abcd2x3</td>
</tr>
<tr>
<td>metricType:</td>
<td>Instantaneous</td>
</tr>
<tr>
<td>metricValue:</td>
<td>48.0 C</td>
</tr>
<tr>
<td>objectType:</td>
<td>CELL</td>
</tr>
</tbody>
</table>

CellCLI> LIST METRICCURRENT WHERE objectType = 'CELLDISK' AND
    metricValue != 0 ATTRIBUTES name,
    metricObjectName, metricValue, collectionTime

CD_IO_By_R_LG   CD_00_abcd2x3    1.9 MB
2009-12-17T15:46:52-08:00
7.4.1.3 Displaying Metric History

Use the LIST METRICHISTORY command to display the metric history for a storage server.

A metric history listing shows a collection of past individual observations of all metric values.

The retention period for metric history files is specified by the `metricHistoryDays` cell attribute. The default retention period is 7 days. You can modify this setting with the `CellCLI ALTER CELL` command.

Example 7-16  Displaying Metric History Values

This example shows various LIST METRICHISTORY commands.

The following command displays observations in the metric history for the `CD_IO_RQ_R_LG` metric when the metric was in a critical alert state. The output contains all of the metric attributes.

```
CellCLI> list metrichistory CD_IO_RQ_R_LG WHERE alertState='critical' detail
```

The following command displays a subset of metric attributes for all metric observations that are associated with cell disks. The output is limited to observations with non-zero metric values that were observed after the specified time.

```
CellCLI> list metrichistory WHERE objectType = 'CELLDISK' AND metricValue != 0 AND collectionTime > '2009-08-12T09:51:07-07:00' attributes name, metricObjectName, metricValue, collectionTime
```
The following command displays all metric observations within a specified time period that are associated with the specified object types.

CellCLI> LIST METRICHISTORY WHERE objectType LIKE 'CELLDISK|FLASHCACHE|FLASHLOG|SMARTIO|IORM_DATABASE' - AND collectionTime > '2020-07-15T08:00:00-07:00' - AND collectionTime < '2020-07-15T09:00:00-07:00'

Related Topics
- DESCRIBE METRICHISTORY
- LIST METRICHISTORY
- ALTER CELL
- DESCRIBE CELL

7.4.2 Monitoring Alerts

You can monitor and receive notifications for alerts. Alerts represent events of importance occurring within the cell. Typically, alerts indicate that Exadata server functionality is compromised or in danger of failure.

To receive notifications, use the ALTER CELL command.

- Displaying Alert Definitions
  Use the LIST ALERTDEFINITION command to display the alert definitions for the storage server.

- Receiving Alert Notifications
  Administrators for Oracle Exadata System Software can receive alert notifications by email or by Simple Network Management Protocol (SNMP) trap alerts.

- Displaying Alert History
  Use the LIST ALERTHISTORY command to display the alert history that has occurred on a cell.

- Modifying Alert History
  Use the ALTER ALERTHISTORY command to update the alert history for the cell.

Related Topics
- Exadata Alerts
  Alerts draw attention to potential and actual problems and other events of interest to an administrator.
- ALTER CELL

7.4.2.1 Displaying Alert Definitions

Use the LIST ALERTDEFINITION command to display the alert definitions for the storage server.

An alert definition provides a definition for every alert that can be produced on a storage server.
### Example 7-17  Listing Alert Definition Attributes

This example shows how to display a detailed list of attributes for the alert `ADRAlert`.

```plaintext
CellCLI> LIST ALERTDEFINITION ADRAlert DETAIL
name:                   ADRAlert
alertShortName:         ADR
alertSource:            "Automatic Diagnostic Repository"
description:            "Incident Alert"
```

### Example 7-18  Listing Alert Definition Name and Description Attributes

You can display a list of specified attributes for an alert definition. This example shows how to display the alert name, metric name, and description. The metric name identifies the metric on which the alert is based. `ADRAlert`, `HardwareAlert`, `Stateful_HardwareAlert`, and `Stateful_SoftwareAlert` are not based on a metric, and therefore do not have metric names.

```plaintext
CellCLI> LIST ALERTDEFINITION ATTRIBUTES name, metricName, description
ADRAlert                                        "Incident Alert"
HardwareAlert                                   "Hardware Alert"
MetricAlert                                      "Threshold Alert"
SoftwareAlert                                    "Software Alert"
StatefulAlert_CD_IO_ERRS_MIN                    CD_IO_ERRS_MIN  "Threshold Alert"
StatefulAlert_CD_IO_RQ_LG                        CG_IO_RQ_LG     "Threshold Alert"
StatefulAlert_CD_IO_RQ_LG_SEC                    CG_IO_RQ_LG_SEC "Threshold Alert"
StatefulAlert_CD_IO_RQ_SM                        CG_IO_RQ_SM     "Threshold Alert"
StatefulAlert_CD_IO_RQ_SM_SEC                    CG_IO_RQ_SM_SEC "Threshold Alert"
StatefulAlert_CT_IO_RQ_LG                        CT_IO_RQ_LG     "Threshold Alert"
StatefulAlert_CT_IO_RQ_LG_SEC                    CT_IO_RQ_LG_SEC "Threshold Alert"
StatefulAlert_CT_IO_RQ_SM                        CT_IO_RQ_SM     "Threshold Alert"
StatefulAlert_CT_IO_RQ_SM_SEC                    CT_IO_RQ_SM_SEC "Threshold Alert"
StatefulAlert_CL_FSUT                            CL_FSUT         "Threshold Alert"
StatefulAlert_CL_MEMUT                           CL_MEMUT        "Threshold Alert"
StatefulAlert_DB_IO_RQ_LG                        DB_IO_RQ_LG     "Threshold Alert"
StatefulAlert_DB_IO_RQ_LG_SEC                    DB_IO_RQ_LG_SEC "Threshold Alert"
StatefulAlert_DB_IO_RQ_SM                        DB_IO_RQ_SM     "Threshold Alert"
StatefulAlert_DB_IO_RQ_SM_SEC                    DB_IO_RQ_SM_SEC "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_LG                     CT_IO_WT_LG     "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_LG_RQ                  CT_IO_WT_LG_RQ  "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_SM                     CT_IO_WT_SM     "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_SM_RQ                  CT_IO_WT_SM_RQ  "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_SEC                    CG_IO_RQ_SM      "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_SEC_RQ                 CG_IO_RQ_SM_SEC "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_WT                    CT_IO_WT_WT     "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_WT_RQ                  CT_IO_WT_WT_RQ  "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_WT_SEC                 CT_IO_WT_WT_SEC "Threshold Alert"
StatefulAlert_DB_IO_RQ_WT_WT_SEC_RQ              CT_IO_WT_WT_SEC_RQ "Threshold Alert"
StatefulAlert_DB_IO_WT_LG                        DB_IO_WT_LG     "Threshold Alert"
StatefulAlert_DB_IO_WT_LG_RQ                     DB_IO_WT_LG_RQ  "Threshold Alert"
StatefulAlert_DB_IO_WT_SM                        DB_IO_WT_SM     "Threshold Alert"
StatefulAlert_DB_IO_WT_SM_RQ                     DB_IO_WT_SM_RQ  "Threshold Alert"
StatefulAlert_DB_IO_WT_SM_SEC                    CG_IO_WT_SM      "Threshold Alert"
StatefulAlert_DB_IO_WT_SM_SEC_RQ                 CG_IO_WT_SM_SEC "Threshold Alert"
StatefulAlert_DB_IO_WT_WT                        DB_IO_WT_WT     "Threshold Alert"
StatefulAlert_DB_IO_WT_WT_RQ                     DB_IO_WT_WT_RQ  "Threshold Alert"
StatefulAlert_DB_IO_WT_WT_SEC                    CG_IO_WT_WT     "Threshold Alert"
StatefulAlert_DB_IO_WT_WT_SEC_RQ                 CG_IO_WT_WT_SEC "Threshold Alert"
StatefulAlert_GD_IO_ERRS_MIN                    GD_IO_ERRS_MIN  "Threshold Alert"
Stateful_HardwareAlert                          "Hardware Stateful Alert"
```
Stateful_SoftwareAlert

Related Topics

- **Displaying Metric Definitions**
  Use the `LIST METRICDEFINITION` command to display the metric definitions for a storage server.

- **DESCRIBE ALERTDEFINITION**

- **LIST ALERTDEFINITION**

### 7.4.2.2 Receiving Alert Notifications

Administrators for Oracle Exadata System Software can receive alert notifications by email or by Simple Network Management Protocol (SNMP) trap alerts.

By using the `ALTER CELL` command, you can set the alert notification attributes to configure Oracle Exadata Storage Server to send notification email messages or SNMP trap alerts.

Use of SNMP alerts allows Oracle Exadata Storage Servers to be monitored by a management application, such as Oracle Enterprise Manager Cloud Control.

**Note:**

The SNMP alerts conform to a MIB (management information base) which is included with each installation of Oracle Exadata System Software. The MIB file on Oracle Exadata Storage Server is available at `/opt/oracle/cell/cellsrv/deploy/config/cell_alert.mib`.

Related Topics

- **ALTER CELL**
- **Alert Notification Attributes**
- **Format of Alert Messages for SNMP Notification**

### 7.4.2.3 Displaying Alert History

Use the `LIST ALERTHISTORY` command to display the alert history that has occurred on a cell.

Alert history entries are retained for a maximum of 100 days. If the number of alert history entries exceeds 500, then the alert history entries are only retained for 7 days. When stateful alerts are cleared, meaning that the underlying metric, hardware or software condition has returned to normal, then the stateful alert is retained either 100 or 7 days, depending on the number of alert history entries. Stateful alerts that are not cleared are retained, regardless of their age.
Example 7-19  Listing Alert History Attributes

This example shows how to display a detailed list of attributes for alert history entries where the severity attribute is set to critical and the examinedBy attribute has not been set.

```
CellCLI> LIST ALERTHISTORY WHERE severity = 'critical' AND examinedBy = '' DETAIL
```

Related Topics

- DESCRIBE ALERTHISTORY
- LIST ALERTHISTORY
- ALTER CELL
- DESCRIBE CELL

7.4.2.4 Modifying Alert History

Use the ALTER ALERTHISTORY command to update the alert history for the cell.

Example 7-20  Altering Alert History Attributes

This example shows how to set the examinedBy attribute to the user ID of the administrator that examined the alert.

```
CellCLI> ALTER ALERTHISTORY 1671443714 examinedBy="jdoe"
```

The examinedBy attribute is the only ALERTHISTORY attribute that can be modified.

Related Topics

- ALTER ALERTHISTORY
- DESCRIBE ALERTHISTORY

7.4.3 Displaying Active Requests

Use the LIST ACTIVEREQUEST command to display the active requests for the storage server.

To view the ACTIVEREQUEST attributes, use the DESCRIBE ACTIVEREQUEST command.

Example 7-21  Listing Active Request Attributes

This example shows how to display a detailed list of attributes for a specified request I/O type.

```
CellCLI> LIST ACTIVEREQUEST WHERE IoType = 'predicate pushing' DETAIL
```

Related Topics

- LIST ACTIVEREQUEST
- DESCRIBE ACTIVEREQUEST
Using the CellCLI Utility

You use the Cell Control Command-Line Interface (CellCLI) utility to manage Oracle Exadata System Software.

CellCLI provides many of the features that are provided with SQL*Plus, including the use of script files.

- **Overview of the CellCLI Utility**
  The Cell Control Command-Line Interface (CellCLI) utility is the command-line administration tool for Oracle Exadata System Software.

- **About CellCLI Administration Commands**
  CellCLI administrative commands do not act directly on objects.

- **About CellCLI Object Commands**
  This topic describes the CellCLI object commands, object types, and object attributes.

- **About CellCLI Object Types**
  There are several Oracle Exadata System Software object types that can be used with CellCLI object commands.

- **About Quoting Object Names**
  If an object name has characters that cause parsing errors enclose the object name in quotes.

- **About CellCLI Object Attributes**
  Each CellCLI object has a set of attributes that are assigned when the object is created or altered.

- **CellCLI Command Reference**
  CellCLI has both administrative and object commands.

**Related Topics**

- **Using the dcli Utility**
  The dcli utility facilitates centralized management across Oracle Exadata Database Machine.

8.1 Overview of the CellCLI Utility

The Cell Control Command-Line Interface (CellCLI) utility is the command-line administration tool for Oracle Exadata System Software.

CellCLI runs on each cell to enable you to manage an individual cell. You use CellCLI to start and stop the cell, to manage cell configuration information, to enable or disable cells, and to manage objects in the cell environment.

The command-line utility is already installed when Oracle Exadata Storage Server is shipped.
• **Starting CellCLI**
  You can start CellCLI from the operating system command line on the cell that you want to manage or remotely from a network-attached client using Secure Shell (SSH).

• **Understanding Command Syntax and Options for CellCLI**
  This topic describes the syntax and command options for CellCLI.

• **Reserved Words**
  If any of the reserved keywords are used as values in commands, then they must be enclosed in quotation marks.

• **CellCLI Command-Line Editing**
  Most of the command editing features of CellCLI are similar to modern shells, such as `bash` and `tcsh`.

• **CellCLI Input and Output Options**
  Oracle Exadata System Software command-line utilities read commands from standard input and write output to standard output.

• **Comments in CellCLI Scripts**
  You can add single-line comments to CellCLI scripts using multiple formats.

• **Line Continuation in CellCLI Commands**
  To continue a long command on to the next line, insert a hyphen (-) at the end of the line.

### 8.1.1 Starting CellCLI

You can start CellCLI from the operating system command line on the cell that you want to manage or remotely from a network-attached client using Secure Shell (SSH).

**Syntax**

The command-line syntax is as follows:

```
```

The `port_number` in the preceding command specifies the HTTP port number of the Management Server (MS) for the cell. If the `port_number` argument is omitted, then the CellCLI utility uses the value assigned to the `HTTP_PORT` variable in the `cellinit.ora` file on the cell. If the `HTTP_PORT` variable in the `cellinit.ora` file is not set, then the port number defaults to 8888.

**Options**

The following options can be used with the CellCLI command:

- `-n` — Runs the CellCLI utility in non-interactive mode. This option suppresses the command prompt and disables the command-line editing features.
- `-m` — Runs CellCLI monitor (read-only) mode.
- `-xml` — Causes output to be displayed in XML format for the Oracle Enterprise Manager.
- `-v`, `-vv`, and `-vvv` — Sets the log level. The `-v` option is for fine, `-vv` is for finer, and `-vvv` is for the finest level.
• –x — Suppresses the banner.

• –e command — Runs the specified CellCLI command. CellCLI exits after running the command. For example:

```
$ cellcli -e list cell detail
$ cellcli -e "list celldisk attributes name where name -like '.*cell01'"
```

Authentication

CellCLI does not have a login parameter or a connect command. CellCLI uses the cell operating system authentication. The directory from which CellCLI is invoked is the default directory for unqualified file access in CellCLI SPOOL and START commands.

8.1.2 Understanding Command Syntax and Options for CellCLI

This topic describes the syntax and command options for CellCLI.

CellCLI syntax is as follows:

```
{admin-command | object-command object} [options] ;
```

In the preceding syntax, the following arguments are used:

• `admin-command` is an administrative action.

• `object-command` is an action performed on an object.

• `object` is an object or target on which a command performs an action.

• `options` extend the use of a command combination to include additional parameters for the command.

When using the CellCLI utility, the following rules apply:

• Commands, objects, and options are not case-sensitive except where explicitly stated, such as in string patterns used in filtering strings with the LIKE operator.

• Use single quotation marks or double quotation marks around the name of an object that includes spaces or punctuation. The use of quotation marks should match. For example, "this is incorrect" is incorrect because the first mark is double quotation marks, and the second is a single quotation mark.

• The current, local cell is the cell to which all CellCLI commands apply.

• A semicolon (;) is optional at the end of a CellCLI command.

• A hyphen (-) is used at the end of a line to continue a long command onto the next line. If you are using hyphens in names or to denote negative values, it must be immediately followed by an alphanumerical value.
8.1.3 Reserved Words

If any of the reserved keywords are used as values in commands, then they must be enclosed in quotation marks.

The following are CellCLI reserved words:

ABORT
ACTIVE
ACTIVEREQUEST
ALERTDEFINITION
ALERTHISTORY
ALL
ALTER
ASSIGN
BBU
BMC
CALIBRATE
CATPLAN
CELL
CONFIGUREBMC
CREATE
DBPLAN
DESCRIBE
DETAIL
DROP
IMPORT
FLASHCACHE
FLASHCACHECONTENT
FOR REPLACEMENT
FORCE
GRIDDISK
IGNORE REDUNDANCY
INACTIVE
IORMPLAN
KEY
LED
LIST
LUN
MAIL
MEMORY
METRICDEFINITION
METRICCURRENT
METRICHISTORY
MS
NULL
ON
PHYSICALDISK
PRIVILEGE
REALM
RESTART
RS
8.1.4 CellCLI Command-Line Editing

Most of the command editing features of CellCLI are similar to modern shells, such as bash and tcsh.

The CellCLI utility supports command-line history and editing, similar to Berkeley Software Distribution (BSD) `editline` and GNU `readline` functionality.

8.1.5 CellCLI Input and Output Options

Oracle Exadata System Software command-line utilities read commands from standard input and write output to standard output.

You can use the host operating system options for redirecting input and output to compose and process command scripts. For example, you can perform the following redirection:

```
$ CellCLI < command-script-in > results-out
```

In this example, the output from CellCLI commands in the `command-script-in` file are written to the `results-out` file.

8.1.6 Comments in CellCLI Scripts

You can add single-line comments to CellCLI scripts using multiple formats.

You can begin the comment line with `REMARK`, `REM` or `--` (two hyphens).

For example, the following are valid syntax for comments:

```
REMARK This is a comment
REM This is a comment
-- This is a comment
```

8.1.7 Line Continuation in CellCLI Commands

To continue a long command on to the next line, insert a hyphen (-) at the end of the line.

After inserting the hyphen, press Enter, and continue typing the command.

For example:

```
CellCLI> LIST CELLDISK WHERE name LIKE 'CD_04.*' -
ATTRIBUTES name, status, comment
```
8.2 About CellCLI Administration Commands

CellCLI administrative commands do not act directly on objects.

The following administration commands are available with CellCLI:

- EXIT
- HELP
- QUIT
- SET
- SPOOL
- START and @

Note:

The celladmin user should be used to run all services on the cell. The cellmonitor user is for monitoring purposes. The cellmonitor user can run the following commands:

- DESCRIBE
- EXIT
- HELP
- LIST
- REMARK
- SET
- START

8.3 About CellCLI Object Commands

This topic describes the CellCLI object commands, object types, and object attributes.

The following CellCLI commands operate on Oracle Exadata System Software objects:

- ALTER
- ASSIGN KEY
- CALIBRATE
- CREATE
- DESCRIBE
- DROP
- EXPORT CELLDISK
- GRANT
8.4 About CellCLI Object Types

There are several Oracle Exadata System Software object types that can be used with CellCLI object commands.

- **ACTIVEREQUEST** — An active request provides a client-centric or application-centric view of client I/O requests that are currently being processed by a cell. The active request object can be used only with the LIST command.
- **ALERTDEFINITION** — An alert definition provides a definition for every alert that can be produced on the cell. Alerts are defined on metrics and other sources of alerts.
- **ALERTHISTORY** — An alert history provides a list of alerts that have occurred on the cell.
- **ASMCLUSTER** — Indicates that the security key is for a client is an Oracle ASM cluster.
- **CELL** — Cell refers to the current or local cell. A cell is the server to which disks are attached and on which the CellCLI utility runs.
- **CELLDISK** — Each cell disk is associated with a logical unit number (LUN). One physical disk is associated with each cell disk.
- **DATABASE** — Database refers to an active database instance.
- **DIAGPACK** — A diagpack represents a compressed file under $LOG_HOME and contains log files and trace files.
- **FLASHCACHE** — The portion of flash storage allocated for use as a cache.
- **FLASHCACHECONTENT** — List of all objects currently cached in the flash cache.
- **FLASHLOG** — The portion of flash storage allocated for storing the Exadata Smart Flash Log.
- **GRIDDISK** — A grid disk is a logical partition of a cell disk. It is exposed on the Oracle Exadata Storage Server network to the database hosts, where it becomes an Oracle ASM disk that can be used to create an Oracle ASM disk group.
- **IBPORT** — The InfiniBand Network Fabric ports for Oracle Exadata Storage Server. Does not apply to Oracle Exadata Database Machine X8M systems.
- **IORMPLAN** — An I/O Resource Management (IORM) interdatabase plan is a set of directives that determines allocation of I/O resources to database clients. There is one plan for the cell.
- **IORMPROFILE** — IORM interdatabase plans support profiles to ease management, and configuration of interdatabase plans for hundreds of databases. Profiles introduce a way to allocate I/O resources for a database.
• **KEY** — A key is a unique hexadecimal string that identifies clients for security purposes.

• **LUN** — Logical unit number (LUN) is the address for an individual physical disk device (a single-disk LUN). LUNs are automatically discovered when the cell is started. They are assigned to the corresponding cell disk when the cell disk is first created or when cell disks are discovered after the system is restarted. LUNs that are not yet assigned to a cell disk have a NULL value for the `cellDisk` attribute.

• **METRICCURRENT** — A current metric describes a set of observations on the current value of an individual metric.

• **METRICDEFINITION** — A metric definition describes the configuration of a metric.

• **METRICCURRENT** — A metric current describes a set of observations on the current value of an individual metric.

• **OFFLOADGROUP** — An object that contains modifiable attributes of offload groups, and can be used to restart, start up, and shut down services.

• **OFFLOADPACKAGE** — An object that contains attributes describing offload support on the cell, including the installation time. This was added starting in version 12.2. Only LIST is supported.

• **PHYSICALDISK** — A disk is called a physical disk on the cell. Physical disks can be listed, but they are not managed directly by CellCLI. Physical disks are automatically discovered and assigned to the corresponding cell disk when the cell disk is first created or when cell disks are discovered after the system is restarted.

• **PLUGGABLEDATABASE** — This object describes tenant databases known to the cell.

• **PMEMCACHE** — The portion of PMEM storage allocated for use as a cache.

• **PMEMLOG** — The portion of PMEM storage allocated for caching redo log data.

• **PRIVILEGE** — A right or permission assigned to a role.

• **QUARANTINE** — A quarantine stops faulty SQL statements from performing a Smart Scan. This reduces software crashes, and improves storage availability.

• **ROLE** — A named group of related privileges.

• **SOFTWAREHISTORY** — A list of final states for past software updates.

• **SOFTWAREUPDATE** — An object that contains the software location and time parameters for scheduling software updates

• **THRESHOLD** — A threshold describes the rules for generating stateful alerts based on a specific metric. The rules include boundary (threshold) values and how long the metric values can violate these boundaries before an alert is generated.

• **USER** — A person allowed access to the storage servers.

Not all possible command-object combinations are valid. For valid command-object combinations, review the syntax for the specific object command.

**Related Topics**

• **About CellCLI Object Commands**
  This topic describes the CellCLI object commands, object types, and object attributes.

• **Oracle Exadata System Software Components**
  This section provides a summary of the following Oracle Exadata System Software components.
8.5 About Quoting Object Names

If an object name has characters that cause parsing errors enclose the object name in quotes.

The object name can be a number or string in commands that have the following format:

\(<\text{verb}>\ <\text{object\_type}>\ <\text{object\_name}>\)

For example, consider the following command:

```
CellCLI> list pluggabledatabase CDB$ROOT
```

CELL-01504: Invalid command syntax.

LIST is the verb, PLUGGABLEDATABASE is the object type, and CDB$ROOT is the object name. In the above example, the parser throws an error when confronted by the $ character in the object name.

For object names that contain invalid characters, you need to surround the object name with either single or double quotes to force the parser to treat the object name as a string.

To avoid this error, surround the object name with either single or double quotes:

```
CellCLI> LIST PLUGGABLEDATABASE 'CDB$ROOT'
```

8.6 About CellCLI Object Attributes

Each CellCLI object has a set of attributes that are assigned when the object is created or altered.

Attribute filters and lists are used to specify which attributes and objects are displayed in the output of the LIST command.

All attributes can be displayed, but only some can be modified directly by the user.
To display a list of attributes and determine which ones can be modified, use the DESCRIBE command.

- Restrictions on Values of Common Attributes
  Review the following restrictions for the values of attributes used by multiple CellCLI objects.

- Attribute Lists in LIST Command
  You can specify which attributes to display for the LIST command with the optional ATTRIBUTES clause.

- Attribute Filters in LIST and ALTER GRIDDISK Commands
  You can specify which objects to display for the LIST and ALTER GRIDDISK commands with the optional attribute_filters clause.
8.6.1 Restrictions on Values of Common Attributes

Review the following restrictions for the values of attributes used by multiple CellCLI objects.

- The value of the **name** attribute must be less than 256 characters and composed only of the following ASCII characters:
  - Lowercase alphabetic characters (a to z)
  - Uppercase alphabetic characters (A to Z)
  - Numbers (0 to 9)
  - Underscore (_)
  - Hyphen (-)

**Note:**

On Oracle Exadata System Software release 19.2.0, or earlier, you must enclose the string in double quotes. For example: "hyphenated-string"

- The value of the **comment** attribute must be less than 256 characters.

See the syntax of each CellCLI command for any additional restrictions on attribute values.

8.6.2 Attribute Lists in LIST Command

You can specify which attributes to display for the **LIST** command with the optional **ATTRIBUTES** clause.

This syntax of the **ATTRIBUTES** clause is:

```
ATTRIBUTES { ALL | attribute1 [, attribute2] ... }
```

**ALL** displays all possible object attributes for the **LIST** object.

**Example 8-1   Listing METRIC HISTORY for Specific Attributes**

This example shows the **LIST METRIC HISTORY** command with the **name** and **metrictype** attributes specified, and the output.

```
LIST METRIC HISTORY ATTRIBUTES name, metrictype
  CL_CPUT           Instantaneous
  CL_FANS           Instantaneous
  CL_RUNQ           Instantaneous
  CL_TEMP           Instantaneous
  N_NIC_RCV_SEC     Rate
```
8.6.3 Attribute Filters in LIST and ALTER GRIDDISK Commands

You can specify which objects to display for the LIST and ALTER GRIDDISK commands with the optional attribute_filters clause.

This syntax of the attribute_filters clause is:

```
WHERE attribute_filter1 [ AND attribute_filter2 ... ]
```

Each attribute_filterN has the following syntax:

```
attribute [ NOT | ! ] operator comparison_value
```

The attribute placeholder represents the name of the attribute to use for filtering. The supported types of operator are listed in the following table. These operators can be combined with NOT or !.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>=</td>
<td>Tests for equality between string, status, or numeric attributes. For example: status NOT = normal</td>
</tr>
<tr>
<td>&gt;</td>
<td>Tests for values greater than the numeric attributes. For example: size &gt; 139920M</td>
</tr>
<tr>
<td>&lt;</td>
<td>Tests for values less than the numeric attributes. For example: freeSpace !&lt; 100M</td>
</tr>
<tr>
<td>LIKE</td>
<td>Tests for a regular expression match with a string attribute using case-sensitive matching. For example: LIKE 'GD_IO_RQ.*'</td>
</tr>
</tbody>
</table>

When used with the supported operators, comparison_value is one of the following data types:

- Numeric
- Literal: Value such as active or normal
- Datetime: Time value supported only for ALERTHISTORY
- String: Value delimited by single quotation marks (') or double quotation marks ("")
- NULL: Unassigned strings or empty lists
8.7 CellCLI Command Reference

CellCLI has both administrative and object commands.

The following commands are available with the CellCLI utility:

- ALTER
- ASSIGN KEY
- CALIBRATE
- CREATE
- DESCRIBE
- DROP
- EXIT
- EXPORT CELLDISK
- GRANT
- HELP
- IMPORT CELLDISK
- LIST
- QUIT
- REVOKE
- SET
- SPOOL
- START and @

8.7.1 ALTER

Purpose

The ALTER command performs an action on or changes attributes of a single cell object or multiple Oracle Exadata System Software objects. The ALTER command can be used to change an attribute or to take an action upon the object.

Syntax

```
ALTER { object_type object_name [, object_name]... operation
  | attribute_name = attribute_value
  [, attribute_name = attribute_value]...
}
```

Usage Notes

When multiple objects are the target of an ALTER command, there is the possibility of partial success. If an error occurs, then the command is interrupted, and the remaining objects are not changed.
ALTER ALERTHISTORY
ALTER CELL
ALTER CELLDISK
ALTER FLASHCACHE
ALTER GRIDDISK
ALTER IBPORT
ALTER IORMPLAN
ALTER LUN
ALTER OFFLOADGROUP
ALTER PHYSICALDISK
ALTER PMEMCACHE
ALTER QUARANTINE
ALTER SOFTWAREUPDATE
ALTER THRESHOLD
ALTER USER

Related Topics

• About CellCLI Object Types
  There are several Oracle Exadata System Software object types that can be used with CellCLI object commands.

8.7.1.1 ALTER ALERTHISTORY

Purpose

The ALTER ALERTHISTORY command changes the attributes of all or specified alert histories.

Syntax

```
ALTER ALERTHISTORY { ALL | alertid1 [,alertid2 ...] } 
examinedBy=user_name
```

Usage Notes

The following arguments can be used with the command:

• `alertidn`: The identifier of the alerts to be changed.
• `user_name`: The name of the user who acknowledged the alert.

Example 8-2  Altering ALERTHISTORY Attributes

This example shows the ALTER command used with the ALERTHISTORY object to update the examinedBy attribute. The examinedBy attribute is the only ALERTHISTORY attribute that can be modified.

```
CellCLI> ALTER ALERTHISTORY 1671443714 -
examinedBy="jdoe"
```
CellCLI> ALTER ALERTHISTORY ALL examinedBy="jdoe"

8.7.1.2 ALTER CELL

Purpose

The ALTER CELL command changes the attributes of the cell.

Syntax

```
ALTER CELL {
| SHUTDOWN SERVICES { RS | MS | CELLSRV | ALL } [IGNORE REDUNDANCY]
| RESTART SERVICES { RS | MS | CELLSRV | ALL } [IGNORE REDUNDANCY]
| LED { ON | OFF }
| DONOTSERVICED { ON | OFF [FORCE] }
| VALIDATE { MAIL | SNMP | CONFIGURATION }
| VALIDATE SYSLOGCONF selector.node
| CONFIGUREBMC
| BBU { DROP FOR REPLACEMENT | REENABLE }
| { snmpuser=((user_clause)[,(user_clause)][,...]) | snmpuser.name=(user_clause) }
| attribute_name = attribute_value
| [, attribute_name = attribute_value]...
}
```

• ALTER CELL Commands for Managing Services
• ALTER CELL Commands for Managing Exadata Storage Server Hardware
• ALTER CELL Commands for Configuration Validation
• ALTER CELL Commands for Setting Attributes

8.7.1.2.1 ALTER CELL Commands for Managing Services

Syntax

```
ALTER CELL {
| SHUTDOWN SERVICES { RS | MS | CELLSRV | ALL } [IGNORE REDUNDANCY]
| RESTART SERVICES { RS | MS | CELLSRV | ALL } [IGNORE REDUNDANCY]
| STARTUP SERVICES { RS | MS | CELLSRV | ALL }
}
```

Usage Notes

The following table lists the arguments and options for the ALTER CELL commands that perform service management operations:
<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHUTDOWN SERVICES {RS</td>
<td>MS}</td>
</tr>
<tr>
<td>SHUTDOWN SERVICES CELLSRV [IGNORE REDUNDANCY]</td>
<td>Shuts down the Cell Server service. If you include IGNORE REDUNDANCY, then the service is immediately stopped without waiting on the redundancy checks from Oracle ASM.</td>
</tr>
<tr>
<td>SHUTDOWN SERVICES ALL</td>
<td>Shuts down all services (Restart Server, Management Server, and Cell Server). If you include IGNORE REDUNDANCY, then the Cell Server service is immediately stopped without waiting on the redundancy checks from Oracle ASM.</td>
</tr>
<tr>
<td>RESTART SERVICES {RS</td>
<td>MS}</td>
</tr>
<tr>
<td>RESTART SERVICES CELLSRV [IGNORE REDUNDANCY]</td>
<td>Stops and then starts the Cell Server service. If you include IGNORE REDUNDANCY, then the service is stopped without waiting on the redundancy checks from Oracle ASM.</td>
</tr>
<tr>
<td>RESTART SERVICES ALL [IGNORE REDUNDANCY]</td>
<td>Stops and then starts all services (Restart Server, Management Server, and Cell Server). If you include IGNORE REDUNDANCY, then the Cell Server service is stopped without waiting on the redundancy checks from Oracle ASM.</td>
</tr>
<tr>
<td>STARTUP SERVICES {RS</td>
<td>MS</td>
</tr>
</tbody>
</table>

The following are additional usage notes for the `ALTER CELL` commands that perform service management:

- During a shutdown operation affecting CELLSRV, the system first checks the status of the grid disks to ensure that it is safe to proceed. Specifically, the `asmDeactivationOutcome` attribute is checked for all of the grid disks. If the attribute value is `yes`, then the grid disk can be deactivated without data loss.

Depending on the command, the following occurs if the `asmDeactivationOutcome` attribute value is `yes` for all of the grid disks:

- For the `ALTER CELL SHUTDOWN SERVICES CELLSRV` and `ALTER CELL SHUTDOWN SERVICES ALL` commands:
  1. The grid disks are deactivated on the cell.
  2. Oracle ASM takes the corresponding ASM disks offline.
  3. Finally, the applicable services are shutdown.

- For the `ALTER CELL RESTART SERVICES CELLSRV` and `ALTER CELL RESTART SERVICES ALL` commands, the CELLSRV service is restarted immediately, followed by a restart of the MS and RS services, if applicable.

Otherwise, if the `asmDeactivationOutcome` attribute value is not `yes` for any grid disk, then the `CELL-01548` error message is displayed and the status of the services remains unchanged.
The **IGNORE REDUNDANCY** option bypasses the `asmDeactivationOutcome` attribute checks. Using the **IGNORE REDUNDANCY** option results in immediate execution of the command. Consequently, if the command shuts down the only online copy of a grid disk, then the corresponding Oracle ASM disk group is dismounted.

- If the Restart Server (RS) service is not running, then you must run either **ALTER CELL STARTUP SERVICES RS** or **ALTER CELL RESTART SERVICES RS** **before you can start other services individually**, or you can run the **ALTER CELL STARTUP ALL** command.

**Example 8-3**  Starting Up and Shutting Down Cell Services

This example shows how to start up and shut down cell services.

```
CellCLI> ALTER CELL STARTUP SERVICES CELLSRV
CellCLI> ALTER CELL STARTUP SERVICES ALL

CellCLI> ALTER CELL SHUTDOWN SERVICES MS
CellCLI> ALTER CELL SHUTDOWN SERVICES CELLSRV IGNORE REDUNDANCY
CellCLI> ALTER CELL SHUTDOWN SERVICES ALL

CellCLI> ALTER CELL RESTART SERVICES ALL IGNORE REDUNDANCY
```

**8.7.1.2.2 ALTER CELL Commands for Managing Exadata Storage Server Hardware**

**Syntax**

```
ALTER CELL {
  RESTART BMC
  | LED {ON | OFF}
  | DONOTSERVICELED {ON | OFF [FORCE]}
  | CONFIGUREBMC
  | BBU { DROP FOR REPLACEMENT | REENABLE }
  | attribute_name = attribute_value [, attribute_name = attribute_value]...
}
```

**Usage Notes**

The following table lists the arguments and options for the **ALTER CELL** commands that perform Exadata storage server hardware management operations:

<table>
<thead>
<tr>
<th>Argument</th>
<th>Options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESTART BMC</td>
<td>none</td>
<td>Restarts the Baseboard Management Controller (BMC).</td>
</tr>
<tr>
<td>Argument</td>
<td>Options</td>
<td>Description</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>LED</td>
<td><strong>ON</strong></td>
<td>LED ON and LED OFF operations turn on and off the Fault-Service Required LED.</td>
</tr>
<tr>
<td></td>
<td><strong>OFF</strong></td>
<td>You can manually light the LED to indicate that a cell requires maintenance.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The LED also turns on automatically if a component fails.</td>
</tr>
<tr>
<td>DONOTSERVICELED</td>
<td><strong>ON</strong></td>
<td>Turns the Do Not Service LED on and off. This LED is available with Oracle Exadata Database Machine X7 and later models.</td>
</tr>
<tr>
<td></td>
<td><strong>OFF</strong></td>
<td></td>
</tr>
<tr>
<td>CONFIGUREBMC</td>
<td>none</td>
<td>Configures the BMC for hardware alerts to the local cell so that Management Server (MS) can pick up the alerts.</td>
</tr>
<tr>
<td>BBU</td>
<td><strong>DROP FOR REPLACEMENT</strong></td>
<td><strong>BBU DROP FOR REPLACEMENT</strong> drops the hard disk controller battery-backed unit (BBU).</td>
</tr>
<tr>
<td></td>
<td><strong>REENABLE</strong></td>
<td><strong>BBU REENABLE</strong> re-enables the BBU.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following are additional usage notes for the `ALTER CELL` commands that perform storage server hardware management:

- **The `ALTER CELL BBU DROP FOR REPLACEMENT` command is run prior to replacement of a hard disk controller battery.** The command changes the caching policy from `writeback` to `writethrough`, and turns on the locator LED. The new battery is enabled automatically.

- **The `ALTER CELL BBU REENABLE` command is run when a battery is removed and then the same battery is re-inserted.** The command changes the caching policy from `writethrough` to `writeback`, and turns off the locator LED.

**Attributes Related to Hardware Management**

- **The `bbuLearnCycleTime` attribute is used to set the start time for the battery learn cycle.** After the learn cycle has completed, the attribute reverts to its default quarterly cycle.

- **The `bbuLearnSchedule` attribute is used to set the next battery learn cycle.** The following parameters are used with the `bbuLearnSchedule` attribute:
  - **month:** Values are 1 through 12. The month entered must be within the current month and the next three months. For example, if the `bbuLearnSchedule` attribute is set in February, then the months could be February, March, April or May.
  - **week:** Values are 1 through 5. The value 1 represents the first week of the month, 2 represents the second week, and so on. The `week` value must be specified when specifying `month` and `day`. 

---

Chapter 8

CellCLI Command Reference

8-17
day: Values are 1 through 7. The value 1 represents Sunday, 2 represents Monday, and so on. The day value must be specified when specifying month and week.

date: Values are 1 through 31. The values represent the days of the month. The default date is 17.

hour: Values are 0 through 23. The value 0 represents 12:00 a.m., 1 represents 1:00 a.m., and so on.

minute: Values are 0 to 59. The values represent the minutes in an hour.

second: Values are 0 to 59. The values represent the seconds in a minute.

- The ALTER CELL eighthRack=true command enables or disables an Eighth Rack configuration on Oracle Exadata Database Machine X3-2 Quarter Racks or later. The options are true to enable the Eighth Rack configuration, and false to disable the Eighth Rack configuration. The ALTER CELL eighthRack=true command requires that there are no cell disks because enabling the Eighth Rack configures only half of the hard disks and flash capacity. After using this command you must restart Cell Server (CELLSRV) to make the new changes effective and prevent unexpected results.

Examples

Example 8-4  Setting the Cell LED Off and On

This example shows how to set the Fault-Service Required LED on and off for the cell.

```
CellCLI> ALTER CELL LED OFF
CellCLI> ALTER CELL LED ON
```

Example 8-5  Setting the Battery Learn Cycle

This example shows how to schedule for the battery learn cycle. In this example, the command sets the battery learn cycle to occur January 17 3:00:59, and then the following learn cycles are April 17 3:00:59, July 17 3:00:59, and October 17 3:00:59. The default setting is "MONTH 1 DATE 17 HOUR 2 MINUTE 0."

```
CellCLI> ALTER CELL bbuLearnSchedule = "MONTH 1 HOUR 3 SECOND 59"
```

8.7.1.2.3 ALTER CELL Commands for Configuration Validation

Syntax

```
ALTER CELL {
  VALIDATE { MAIL | SNMP | CONFIGURATION }
  | VALIDATE SYSLOGCONF selector.node
}
```

Usage Notes

The following table lists the arguments and options for the ALTER CELL commands that perform configuration validation operations:
<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VALIDATE MAIL</td>
<td>The VALIDATE MAIL operation sends a test message using the e-mail attributes configured for the cell.</td>
</tr>
<tr>
<td>VALIDATE SNMP</td>
<td>The VALIDATE SNMP operation sends a test message using the SNMP attributes configured for the cell. The VALIDATE SNMP TYPE=ASR operation validates Oracle ASR on Oracle Exadata Storage Server.</td>
</tr>
<tr>
<td>VALIDATE CONFIGURATION</td>
<td>The VALIDATE CONFIGURATION operation validates the configuration. When the validation is complete and correct, the system responds with Cell cell_name successfully altered. If there is a problem, then the system responds with an error message.</td>
</tr>
<tr>
<td>VALIDATE SYSLOGCONF</td>
<td>The VALIDATE SYSLOGCONF facility.priority sends a test message for the specified facility and priority.</td>
</tr>
</tbody>
</table>

Usage Notes
For more information about SYSLOG configuration, see SYSLOG Attributes.

Examples

Example 8-6 shows how to validate the e-mail setup on a cell.

Example 8-7 shows how to validate the Oracle ASR e-mail setup on a cell.

Example 8-8 shows how to validate the SNMP setup on a cell.

Example 8-9 shows how to validate the configuration on a cell.

Example 8-10 shows a sample error message when configuration on a cell is incorrect.

Example 8-6  Validating E-mail on a Cell

CellCLI> ALTER CELL VALIDATE MAIL

Example 8-7  Validating Oracle ASR E-mail on a Cell

CellCLI> ALTER CELL VALIDATE SNMP type=asr

Example 8-8  Validating SNMP on a Cell

CellCLI> ALTER CELL VALIDATE SNMP
Example 8-9  Validating Configuration on a Cell

CellCLI> ALTER CELL VALIDATE CONFIGURATION

Cell CD_01_cell01 successfully altered

Example 8-10  Checking an Incorrect Configuration on a Cell

CellCLI> ALTER CELL VALIDATE CONFIGURATION

CELL-02827: Cell configuration check for hardware and firmware encountered the following issues:

ILOM check has detected the following issue(s):
  Attribute Name : ILOMVersion
  Required       : 3.0.6.10.a r49240
  Found          : 3.0.6.10.a r49385

8.7.1.2.4 ALTER CELL Commands for Setting Attributes

Syntax

ALTER CELL
  attribute_name  = attribute_value
  [, attribute_name  = attribute_value]...

Usage Notes

The attributes that can be changed using the ALTER CELL command are shown as modifiable in Example 8-92 or described below.

- Caching Attributes
- Alert Notification Attributes
- Alert Summary Attributes
- SYSLOG Attributes
- Disk Scrubbing Attributes
- Security Certificate Attributes
- Miscellaneous Attributes

8.7.1.2.4.1 Caching Attributes

Flash Cache Attributes

The flashCacheMode attribute is used to display and set the current value for flash cache. The values are writethrough (the default) or writeback. Note the following about the flashCacheMode attribute:

- If the attribute is modified from writeback to writethrough and there is existing flash cache, then an error is displayed. The flash cache must be flushed and dropped before changing the attribute to writethrough.
If the attribute is to be modified from writethrough to writeback, then flash cache must be dropped before modifying the attribute.

Write back caching can be disabled on the grid disks that do not need caching, such as the grid disks in the RECO disk group. This allows other objects to use the cache space.

See Also:

- ALTER GRIDDISK for information about disabling caching on grid disks
- Oracle Exadata Database Machine Maintenance Guide for information about enabling and disabling flash cache

Flash cache compression is available in Oracle Exadata System Software releases 11.2.3.3.0 and later running on only Oracle Exadata Database Machine X3 and X4 storage servers.

The ALTER CELL flashCacheCompress command enables or disables flash cache compression. The options are true to enable flash cache compression, and false to disable flash cache compression. To enable flash cache compression on Oracle Exadata Database Machine X3 and X4 storage servers, use the following command:

CellCLI> ALTER CELL flashCacheCompress=true

Note:

Oracle Advanced Compression Option is required to enable flash cache compression.

See Also:

Oracle Exadata Database Machine Maintenance Guide for additional information about enabling flash cache compression

Flash cache compression must be disabled before a storage server is downgraded to an earlier release.

RAM Cache Attributes

The ramCacheMode attribute can be set to on, off, or auto. The default value is auto, which means the RAM Cache feature is not enabled. When you modify this attribute, you must restart CELLSRV for the change to take effect.

Examples

Example 8-11 shows how to set the flash cache mode.

Example 8-12 shows how to enable flash cache compression on a cell in Oracle Exadata Database Machine X4-2.
Setting the Mode for Flash Cache

```
CellCLI> ALTER CELL flashcachemode = writeback
```

Enabling Flash Cache Compression

This example shows how to enable flash cache compression for a storage server.

```
CellCLI> ALTER CELL flashCacheCompress=true
```

8.7.1.2.4.2 Alert Notification Attributes

Configuring Alert Notifications

To set up the cell to send notifications about alerts, you can configure the following cell attributes:

- `smtpServer`: Domain name or IP address of the email server used to send alert notifications
- `smtpPort`: Email server port used to send alert notifications
- `smtpUseSSL`: Specification to use Secure Socket Layer (SSL) encryption for alert notifications.
- `smtpFrom`: User name that appears in the From: header of the alert notifications.
- `smtpFromAddr`: Email address that appears in the From: header of the alert notifications. This email address is not authenticated with the email server.
- `smtpToAddr`: Address to which email is sent. It can be a comma-delimited list in quotation marks to allow multiple subscribers to alerts.
- `snmpSubscriber`: List of hosts that subscribe to the SNMP alert notifications
- `snmpUser`: Defines the user who receives SNMP alerts
- `snmpEngineID`: An identifier used by the SNMP managers to subscribe to alerts from the storage cells
- `notificationMethod`: Notification method for alerts
- `notificationPolicy`: Indicator for severity alerts to be sent to subscribers
- `emailFormat`: File format for email messages
- `emailSubscriber`: List of names that subscribe to the alert notifications

The `smtpPort` attribute can be reset to the default value by setting it to an empty string enclosed by quotation marks (`smtpPort=''`).

The `smtpUseSSL` attribute enables Secure Socket Layer (SSL) encryption on the email notifications when the attribute is set to `true`.

The `smtpToAddr` attribute can be used to set a list of comma-delimited email addresses that are the recipients of the alert notification. The list must be enclosed in quotation marks.
The `snmpSubscriber` attribute can be set to a list of SNMP targets to which the SNMP alert notification is sent. These targets are specified as follows:

```
snmpSubscriber=(
    (host=host[,port=port][,community=community][,type=user_type]
    [,fromIP="ip"][,asrmPort="ASRManager_port"])
    ,(host=host[,port=port][,community=community][,type=user_type]
    [,fromIP="ip"][,asrmPort="ASRManager_port"])) ...
```

**Note:**

- The ALTER CELL `... snmpSubscriber` command replaces the existing value set for `snmpSubscriber`. To add to the list of SNMP targets ensure that you include the existing values in the command. The following message is displayed after running the command:

  snmpSubscriber <old_value> has been replaced with <new_value>

  For example:

  snmpSubscriber ((host=hosta)) has been replaced with ((host=hostb))

The `snmpSubscriber` attribute uses the following values:

- The host must be specified as either a host name or an IP address. Enclose the host name or IP address in quotation marks if it contains non-alphanumeric characters.
- The default value for port is 162. This value is optional.
- The default value for community is public. This value is optional.
- The `snmpSubscriber` types are ASR, v3, and v3ASR. Specifying a type value is optional.
  - The default value for type is NULL.
  - The ALTER CELL `snmpSubscriber...type=asr` command sets the Oracle ASR destination for Oracle Exadata Storage Server, and its ILOM. Removing all `snmpSubscriber` entries with type=asr from the SNMP subscriber list disables the trap mechanism for Oracle Exadata Storage Server and its ILOM.
  - The `snmpSubscriber` with type=asr or type=v3ASR should only be configured to point to Oracle ASR Manager.
  - For types v3 and v3ASR, a `snmpUser` must be defined, and the user name is provided instead of community.
  - For the v3ASR type, the user must be defined with `authProtocol = SHA`, and `privProtocol = AES`. These are the only protocols supported by Oracle ASR Manager. Setting the `snmpSubscriber` as type v3ASR also sets the ILOM properties and rules for traps sent by ILOM.
– If type is not specified, then the default is version 1, cell_alert traps. There is no string to specify this type. To use this type, just omit the type field.

• The fromIP field enables you to specify an IP address from which the trap is sent. If this field is not specified, it defaults to the IP address associated with eth0. Use this field if the default IP address is not registered with Oracle ASR Manager. Oracle ASR Manager only processes SNMP traps that are sent from IP addresses that it recognizes.

The fromIP field is allowed only for SNMP subscribers whose type is either ASR or v3ASR.

For example:

CellCLI> alter cell
snmpSubscriber=((host=asrhost,port=162,community=public,fromIP="1.1.1.1",type=ASR))

The following example returns an error because the type is not ASR or v3ASR.

CellCLI> alter cell
snmpSubscriber=((host=localhost,port=162,community=public,fromIP="1.1.1.1"))

CELL-00068: The fromIP field is only supported for ASR SNMP subscribers.

• The asrmPort field enables you to specify the port number on an Oracle ASR Manager machine that MS uses to communicate with Oracle ASR Manager. This port must be the same as the HTTP port of Oracle ASR Manager's HTTP Receiver. You can check this by running asr show_http_receiver on the Oracle ASR Manager machine.

The asrmPort field is allowed only for SNMP subscribers whose type is either ASR or v3ASR. The default value for this port is 16161.

After startup of the Management Server (MS), the snmpSubscriber list entries with type=asr are added to the ILOM for the cell. This ensures that when an ILOM is replaced, the entries are set for the new ILOM. If the entries are removed from the ILOM, then they must be manually added to the ILOM using the ALTER CELL ... snmpUser= command.

The snmpUser attribute defines the user who receives SNMP alerts. This command can only be run in interactive mode. There are two methods for configuring this attribute.

snmpuser=((user_clause)[,(user_clause)[,..]])

snmpuser.name=(user_clause)

• If you specify snmpuser, then you must provide a user_clause for every configured user. If you omit a user, then that user will no longer receive SNMP alerts. The ((user_clause)[,(user_clause)[,..]]) string that you provide overwrites the previous string used for the snmpuser attribute.
• If you specify snmpuser.name, then you must provide a user_clause for only the specified user. This allows you to add, delete, or modify each user individually, without having to supply the entire snmpuser attribute string each time.

• If you use snmpuser=', then all SNMP users are removed. If you use snmpuser.name=', then only the specified user is removed. You cannot remove an SNMP user while it is still referenced by a V3 SnmpSubscriber.

Each method uses a user_clause, which has the following format:

```plaintext
((name=user1, authProtocol=auth_type, authPassword=*, \\
privProtocol=priv_type, privPassword=*) , \\
(name=user2, authProtocol=auth_type, authPassword=*, \\
privProtocol=priv_type, privPassword=*, ), ...) 
```

If updating a single user, then do not include the phrase name=user1, in the user_clause, because you have already supplied the name as part of snmpuser.name.

• name is the user name.

• Only * is allowed for the password values in the command. Passwords are not stored or displayed. Secure hash keys are computed and used for trap authentication and encryption.

• authProtocol is the authentication protocol. Options are MD5 or SHA.

  The authProtocol must be specified for the snmpUser attribute.

  The system prompts for the authentication password. The authentication password must have 8 to 12 alphanumeric characters.

• privProtocol is encryption protocol. Options are none, AES, or DES. The default is none when the privProtocol attribute is not specified.

  The system prompts for an encryption password if the encryption protocol is specified. The password is exactly 8 alphanumeric characters, and they are case sensitive.

The ALTER CELL snmpEngineID command is used by the SNMP managers to subscribe to alerts from the storage cells. The snmpEngineID parameter can be up to 20 characters. It should be unique for each target within a data center. The default is the cell name. This default is used if the snmpEngineID attribute is not set before the SNMP users are defined.

**Note:**

The engine identifier should not be changed after SNMP users are defined. Any change to an engine identifier causes the user keys to be re-computed, and user passwords must be re-entered.

The notificationMethod attribute value can be mail, snmp, none, or a combination of mail and snmp, such as notificationMethod='mail,snmp'. The default value is mail.

The notificationPolicy attribute value can be none or a combination of critical, warning, or clear, such as notificationPolicy='warning,clear.'
• The **critical** value refers to hardware-generated alerts or alerts generated by Automatic Diagnostic Repository (ADR) or BMC. The **critical** value also refers to a metric alert when the value exceeds the critical threshold specified in the metric definition.

• The **warning** value refers to a metric alert when the value exceeds the warning threshold specified in the metric definition.

• The **clear** value refers to a metric alert when the value is below the threshold boundary after having previously exceeded a warning or critical threshold.

• The **maintenance** value refers to all hardware-related errors. The hardware errors are reported as “Maintenance” in email message subject lines.

To validate that email messages are successfully sent for cell alerts or events, use the ALTER command with the VALIDATE MAIL option. The validation process sends a test email message to the configured recipient. If that test email message is not received, then an email configuration setting is not valid.

The emailFormat attribute can be html or text. By default, email notifications are sent in HTML format. Change the value to text to receive plain text email notifications.

The ALTER CELL emailSubscriber command sets a list of comma-delimited email addresses that are the recipients of alert notifications for specific alert types. The syntax for this command is:

```
ALTER CELL emailSubscriber = ((email="email_address1", alertType="alert_type")
    [, (email="email_address2",alertType="alert_type"), ...])
```

• The email address must be a valid email address. The **email** parameter is mandatory.

• The **alertType** parameter specifies the type of alert, and is optional. The alert types are HARDWARE, SOFTWARE, METRIC or ADR. If the alert type is not specified, then the subscription is for all alert types.

• An empty input string removes the current set of subscribers.

• The notification policy must be set before alert notifications can be received. The policy applies to all email subscribers. The notification policy for these alerts are the same as for snmpSubscriber alerts.

**Examples**

**Example 8-13** shows how to set the asrmPort field for an snmpSubscriber.

**Example 8-14** shows how to set up email notifications for the cell.

**Example 8-15** shows how to specify the type of email alerts. In the example, one subscriber gets hardware and software alerts, and the other subscriber gets ADR alerts.

**Example 8-16** shows how to change the format of email messages.

**Example 8-17** shows how to modify the SNMP user.

**Example 8-18** shows how to modify a single SNMP user.

**Example 8-19** shows how to unsubscribe from email alerts.
Example 8-20 shows how to reset the `notificationPolicy` attribute to its default value.

**Example 8-13  Setting the asrmPort for an snmpSubscriber**

CellCLI> ALTER CELL
snmpSubscriber=((host=host1, port=162, community=public, type=asr, asrmPort=16161))

**Example 8-14  Configuring Email Notifications for a Cell**

```
CellCLI> ALTER CELL smtpServer='my_mail.example.com',
     smtpFromAddr='john.doe@example.com',
     smtpFrom='John Doe',
     smtpToAddr='jane.smith@example.com',
     snmpSubscriber=((host=host1), (host=host2)),
     notificationPolicy='clear',
     notificationMethod='mail,snmp'
```

**Example 8-15  Specifying the Type of Email Alert**

```python
ALTER CELL emailSubscriber=
\"((email="email1@example.com", alertType="HARDWARE,SOFTWARE"),
  (email="email2@example.com", alertType="ADR"))\"
```

**Example 8-16  Changing the Format of Email Messages**

CellCLI> ALTER CELL emailFormat='text'
CellCLI> ALTER CELL emailFormat='html'

**Example 8-17  Modifying the SNMP User**

This example shows the initial configuration of a single SNMP user, where the administrator is prompted to enter the passwords.

```
CellCLI> ALTER CELL snmpuser = ((name=ASR, authprotocol=md5, authpassword=*,
   privprotocol=AES, privpassword=*))

snmpUser ASR authpassword: password
Confirm snmpUser ASR authpassword: password
snmpUser ASR privpassword: password
Confirm snmpUser ASR privpassword: password
```

**Example 8-18  Modifying a Single SNMP User**

The following code examples show adding an SNMP user, changing that user's password, and then removing that user.

```bash
## adding users individually
CellCLI> ALTER CELL snmpuser.user2=(authprotocol=SHA, authpassword=*)
```
Example 8-19  Unsubscribing from Email Alerts

ALTER CELL emailSubscriber=

Example 8-20  Setting the Default Value for the notificationPolicy Attribute

This example shows how to set the default value for the notificationPolicy attribute.

CellCLI> alter cell notificationPolicy=""
Example 8-22  Setting the Time for Open Alerts Message Delivery

This example shows how to set the delivery time for the open alerts summary e-mail message.

CellCLI> ALTER CELL alertSummaryStartTime="2013-04-23T12:57:00-06:00"

8.7.1.2.4.4 SYSLOG Attributes

Configuring SYSLOG Attributes

The syslogconf attribute extends syslog rules for a cell. The attribute can be used to designate that syslog messages be forwarded to a specified management server. On the management server, the forwarded messages are directed to a file, console, or management application, depending on the syslog configuration on the management server. The syntax for configuring this attribute is:

\[
syslogconf = (\text{'selector @node'} [, \text{'selector @node'}] \ldots)
\]

In the preceding syntax, selector is the message type, and node is the specified server. Both variables follow syslog.conf standard syntax rules.

- The facility option for the syslogconf attribute must be one of the following: auth, authpriv, cron, daemon, ftp, kern, lpr, mail, mark, news, security, syslog, user, uucp, local0, local1, local2, local3, local4, local5, local6, local7, none, and *.
- The priority option for the syslogconf attribute must be one of the following: alert, crit, debug, emerg, err, error, info, notice, panic, warn, warning, none, and * (asterisk).

The ALTER CELL VALIDATE syslogconf selector command sends a test log message. The test message is directed as specified by rules in the /etc/syslog.conf file. If the syslogconf assignment extends the syslog rules, then a test message is forwarded to the specified management servers.

Starting with Oracle Exadata System Software release 19.1.0, you can use the syslogFormat attribute to change the standard format for syslog to any format by setting the value to the desired format string. Setting the syslogFormat attribute to an empty string removes the format change, reverting the syslog format to the default format. If the format string contains a control character, it must be preceded by a backslash when entering the command.

See Example 8-26 for examples of the syntax.

Starting with Oracle Exadata System Software release 19.3.0, you can use the syslogFormat attribute to enable sending syslog in an encrypted format. For the complete configuration steps, refer to Encrypting System Log Information.

Example 8-23  Using the syslogconf Attribute

This example shows how to add a rule using the syslogconf attribute.

CellCLI> ALTER CELL syslogconf="*.err;authpriv.none @loghost", -"*.emerg @loghost"
Example 8-24  Adding and Validating a Rule

This example shows how to add and validate a rule with test message.

```
CellCLI> ALTER CELL syslogconf=('kern.crit @loghost')
CellCLI> ALTER CELL VALIDATE syslogconf 'kern.crit'
```

Example 8-25  Removing All syslog.conf Rules

This example shows how to remove the syslog.conf rule.

```
CellCLI> ALTER CELL syslogconf=''
```

Example 8-26  Setting the Syslog Format to a Custom String Then Reverting to the Default Format

This example shows how to specify a customized format for syslog.

```
CellCLI> ALTER CELL syslogformat="%TIMESTAMP:::date-rfc3339% %HOSTNAME% %syslogtag% %syslogseverity-text%:%msg:::sp-if-no-1st-sp%msg:::drop-last-lf%%msg:::drop-last-lf"
CellCLI> ALTER CELL syslogformat="%TIMESTAMP% %HOSTNAME% %msg%"
CellCLI> ALTER CELL syslogformat=""
```

8.7.1.2.4.5 Disk Scrubbing Attributes

Configuring Disk Scrubbing Attributes

The `ALTER CELL hardDiskScrubStartTime` command sets the start time for proactive resilvering of latent bad sectors. Valid options are a date/time combination or `now`. The following table shows the approximate time needed to scrub one idle hard disk:

<table>
<thead>
<tr>
<th>Disk Type</th>
<th>Hard Disk Capacity</th>
<th>Approximate Time for Scrubbing</th>
</tr>
</thead>
<tbody>
<tr>
<td>High performance</td>
<td>600 GB</td>
<td>1 hour</td>
</tr>
<tr>
<td>High performance</td>
<td>1.2 TB</td>
<td>2 hours</td>
</tr>
<tr>
<td>High capacity</td>
<td>2 TB</td>
<td>4 hours, 40 minutes</td>
</tr>
<tr>
<td>High capacity</td>
<td>3 TB</td>
<td>6 hours, 30 minutes</td>
</tr>
<tr>
<td>High capacity</td>
<td>4 TB</td>
<td>8 hours</td>
</tr>
<tr>
<td>High capacity</td>
<td>8 TB</td>
<td>13 hours</td>
</tr>
<tr>
<td>High capacity</td>
<td>10 TB</td>
<td>14 hours</td>
</tr>
<tr>
<td>High capacity</td>
<td>14 TB</td>
<td>18 hours</td>
</tr>
</tbody>
</table>

The `ALTER CELL hardDiskScrubInterval` command sets the interval for proactive resilvering of latent bad sectors. Valid options are `daily`, `weekly`, `biweekly` and `none`. Using the `none` option stops all disk scrubbing.
In the alert log, you may see messages such as Begin scrubbing celldisk and Finished scrubbing celldisk. These informational messages are expected, and no action is necessary.

If the system is idle, disk scrubbing can drive the disk utilization to 100%. This is expected. Disk scrubbing will throttle itself using IORM based on disk activity. When requests come in, disk scrubbing activity will decrease. Your workloads should not be affected by disk scrubbing.

**Example 8-27 Setting the Start Time for Proactive Resilvering of Latent Bad Sectors**

This example shows how to set the start time for resilvering the disks of a storage server.

CellCLI> ALTER CELL hardDiskScrubStartTime='2013-08-07T21:19:22-07:00'

**Example 8-28 Setting the Proactive Resilvering of Latent Bad Sectors Interval to be Weekly**

This example shows how to modify the interval for proactive resilvering of latent bad sectors to weekly.

CellCLI> ALTER CELL hardDiskScrubInterval=weekly

### 8.7.1.2.4.6 Security Certificate Attributes

**Configuring CA-Certified Security Certificate Attributes**

To set up CA-certified security certificates on the cell for use with ExaCLI, use the following attributes:

- **securityPubKey** - Specifies the URL to the public key file.
- **securityPrivKey** - Specifies the URL to the private key file.
- **securityPrivKeyPW** - Specifies the password to use if the private key file is encrypted.

After you upload the CA-certified security certificate, you must restart MS before the new security certificate is visible.

CellCLI> ALTER CELL RESTART SERVICES MS

---

**Note:**

The following attributes can be used only if you are running the ALTER CELL command from ExaCLI.
### Example 8-29 Configuring the Security Keys for a Storage Server

This example shows how to configure the security keys for a storage server, including supplying the password after entering the command.

```bash
exacli -e 'ALTER CELL securityPubKey="http://www.example.com/security/newkey.crt", -
securityPrivKey="http://www.example.com/security/newkey.key", -
securityPrivKeyPW="*
password=*************
```

---

#### 8.7.1.2.4.7 Miscellaneous Attributes

**dbPerfDataSuppress**

Use the `dbPerfDataSuppress` attribute to hide performance output information for specific databases. Specify the databases to mask as a comma-delimited list of names. The performance information for the specified databases is still collected, but is only visible when queried from that database. If you query `V$CELL_DB` from a different database, then the performance information for the hidden databases appears in the category of `OTHER`.

**diagPackEmailAttach**

Use the `diagPackEmailAttach` attribute to turn on and off adding the diagnostic pack attachment to emails, for example:

```bash
alter cell diagPackEmailAttach=FALSE
```

See [CREATE DIAGPACK](#) for information about diagnostic packs.

**diagPackUploadEnabled**

Use the `diagPackUploadEnabled` attribute to enable or disable automatically uploading diagnostic data to a service request using Oracle ASR.

#### Example 8-30 Enabling/Disabling Auto Diagpack Upload

You can enable or disable this feature by setting the `diagPackUploadEnabled` attribute on the `cell` object.
Set the attribute to false to disable this feature, true to enable it. The default is true.

CellCLI> ALTER CELL diagPackUploadEnabled=FALSE

enableSmartStorage

The enableSmartStorage attribute can be set to TRUE to enable the use of Oracle Exadata System Software capabilities such as Smart Scan and Storage Index on Oracle Exadata Storage Server X8-2 Extended (XT) after you have procured the necessary software licenses.

httpsAccess

Starting with Oracle Exadata System Software release 19.1.0, the httpsAccess attribute can be used to specify a list of IP addresses or IP subnet masks that control who can access the RESTful service via HTTPS. The value you specify for httpsAccess overwrites any previous value. You can use the following values for httpsAccess:

- ALL — to allow access to all hosts (Default)
- NONE — to disable the HTTPs port completely
- IP1, IP2,..., IPn — to only allow access to hosts with IP addresses IP1, IP2,..., IPn where IPn is a valid IP address in IPv4, IPv4 subnet, IPv6 or IPv4-embedded IPv6 format. You can specify a maximum of 512 IP addresses for the access control list.

Additionally, instead of a single IP address, you can use the / character to specify a range of IP addresses using a subnet mask. For example the range '192.168.10.0/24' corresponds to hosts having IP addresses from 192.168.10.1 to 192.168.10.255. If you specify an IP address range, you need to enclose the IP address string in quotes.

Example 8-31  Restricting HTTPS Access to the Exadata RESTful Service

This example shows how to configure an access control list for HTTPs access to the Exadata RESTful service. The following command allows HTTPs port access to hosts having IP addresses in the range from 192.168.10.1 to 192.168.10.255.

CellCLI> ALTER CELL httpsAccess="192.168.10.0/24"

interconnectN

The ALTER CELL interconnectN="" command removes the RDMA Network Fabric configuration information for the cell for the specified interface (N).

If the IP address to an RDMA Network Fabric interface is changed, then the command service openibd restart must be run as the root user before the service network restart command.

After changing an IP address, you must restart all services using the ALTER CELL RESTART SERVICES ALL command.

Example 8-32  Setting RDMA Network Fabric Interconnections

This example shows how to set the RDMA Network Fabric interconnections.
For systems that use InfiniBand Network Fabric, use a command such as the following:

CellCLI> ALTER CELL interconnect1='ib0', interconnect2='ib1'

For systems that use RoCE Network Fabric, use a command such as the following:

CellCLI> ALTER CELL interconnect1='re0', interconnect2='re1'

After making the updates, restart all services in the storage server:

CellCLI> ALTER CELL RESTART SERVICES ALL

**iotimeoutthreshold**

Use the `iotimeoutthreshold` attribute to change the timeout threshold. If cell I/O takes longer than the defined threshold, then the I/O is canceled, and Oracle ASM redirects the I/O to another mirror copy of the data. Any I/Os issued to the last valid mirror copy of the data are not canceled, even if the timeout threshold is exceeded.

The default value for `iotimeoutthreshold` is 1000s. The command takes a value, such as 5, and a unit. The valid unit is `s`, for seconds.

**Caution:**

Setting the timeout threshold too low can negatively impact system performance. Oracle recommends reviewing the Automatic Workload Repository (AWR) reports of peak I/O loads, and setting the threshold value to a value higher than the peak I/O latency with sufficient safety margin.

---

**Example 8-33 Setting the iotimeoutthreshold Value**

This example demonstrates how to set the `iotimeoutthreshold` to 5 seconds.

CellCLI> ALTER CELL iotimeoutthreshold = '5s'

To reset the `iotimeoutthreshold` to the default value, use the following command:

CellCLI> ALTER CELL iotimeoutthreshold = ""

**name**

The `name` attribute contains the host name of the storage server, for example, `dm01celladm01`.

**Example 8-34 Altering Cell Name**

This example shows the `ALTER` command with the `CELL` object.

CellCLI> ALTER CELL name=cell02
traceLevel
The level for which trace messages are written. The default is FINE. The value can be:

- A valid Java logging level
  - SEVERE
  - WARNING
  - INFO
  - CONFIG
  - FINE
  - FINER
  - FINEST

- A valid Oracle Diagnostic Logging (ODL) logging level
  - INCIDENT_ERROR:1
  - ERROR:1
  - WARNING:1
  - NOTIFICATION:1
  - NOTIFICATION:16
  - TRACE:1
  - TRACE:16
  - TRACE:32

To reset this attribute to its default value, use a value of "".

Examples

Example 8-35 Setting the traceLevel Value to its Default Value
This example shows how to set the traceLevel value to its default value.

CellCLI> ALTER CELL traceLevel=""

8.7.1.3 ALTER CELLDISK

Purpose
The ALTER CELLDISK command changes the attributes of all cell disks or the specified cell disks.

Syntax

ALTER CELLDISK { ALL [FLASHDISK | HARDDISK | PMEM] | cdisk_name [, cdisk_name]... } {{FLUSH [NOWAIT] | CANCEL FLUSH} | { attribute_name = attribute_value}}
[, attribute_name = attribute_value]...
}

Usage Notes

The attributes that can be changed with the ALTER command are shown as modifiable in Example 8-93.

- The FLASHDISK option limits the ALTER CELLDISK command to cell disks that are flash disks.
- The HARDDISK option limits the ALTER CELLDISK command to cell disks that are hard disks.
- The PMEM option limits the ALTER CELLDISK command to all cell disks of type PMEM.
- The FLUSH option synchronizes dirty data from the flash cache and PMEM cache to the grid disks on the specified cell disks. Dirty data is data that has not been synchronized with the grid disk. Synchronization of dirty data can be a lengthy process, depending on the number of bytes to be synchronized. Use the following command to check the progress:

```sql
LIST CELLDISK ATTRIBUTES name, flushstatus, flusherror
```

- The ALTER CELLDISK ... FLUSH command must be run before exporting a cell disk to ensure that the dirty data is flushed from flash cache and PMEM cache to the grid disks on the specified cell disk.
- The FLUSH option stops new data from being cached on the flash cache until CELLSRV restarts, or the flush operation is canceled.
- The CANCEL FLUSH option terminates an earlier flush operation, and reinstates caching.
- When the ALTER CELLDISK ... FLUSH command is run for a flash-based cell disk, it synchronizes dirty data from the flash cache located on the specified FDOM to cached grid disks on the specified cell disk. When the command is run for a hard disk-based cell disk, it synchronizes dirty data from the flash cache located on all FDOMs to the grid disks located on the specified cell disk.

Example 8-36  Altering Cell Disk Attributes

This example shows how to change cell disk attributes.

```sql
CellCLI> ALTER CELLDISK cdiska name = CD_01_cell01, -
       comment = 'cdiska is now CD_01_cell01'

CellCLI> ALTER CELLDISK ALL -
       comment = 'This cell disk is on cell cell01'

CellCLI> ALTER CELLDISK ALL HARDDISK FLUSH NOWAIT

CellCLI > ALTER CELLDISK c9datafile1 CANCEL FLUSH
```
8.7.1.4 ALTER FLASHCACHE

Purpose
The ALTER FLASHCACHE command stops new data from being cached on the flash cache and then flushes data not synchronized with the grid disks (dirty data) from flash cache to the specified disks.

Syntax
ALTER FLASHCACHE \{ ALL [size=fc_size] | CELLDISK="cdisk1 [,cdisk2] ..." [, size=fc_size] [FORCE] \} { FLUSH [NOWAIT] | CANCEL FLUSH }

Usage Notes

Note:
The \texttt{FLUSH} option stops new data from being cached on the flash cache until CELLSRV restarts, or the flush operation is canceled with the \texttt{ALTER FLASHCACHE CANCEL FLUSH} command.

- The \texttt{ALL} option affects all available flash cell disks.
- The \texttt{CELLDISK} option allows specific cell disks to be flushed.
- The \texttt{FORCE} option allows you to forcefully change the set of cell disks used by the flash cache.
- If specified, the \texttt{size} attribute re-sizes the flash cache. The value is validated.
  
  If a valid size is specified in conjunction with the \texttt{ALL} option, then the flash cache is dropped and re-created on all of the cell disks using the specified size.
  
  If a valid size is specified in conjunction with the \texttt{CELLDISK} option, then the flash cache is dropped and re-created on the specified cell disks using the specified size.
- The \texttt{FLUSH} option synchronizes dirty data from the flash cache to the grid disks. Dirty data is data that has not been synchronized with the grid disk. Synchronization of dirty data can be a lengthy process, depending on the number of bytes to be synchronized. Use the following command to check the progress:

  \begin{verbatim}
  LIST CELLDISK ATTRIBUTES name, flushstatus, flusherror
  \end{verbatim}

- The \texttt{ALTER FLASHCACHE CELLDISK= ... FLUSH} command does not flush the dirty data when the data cannot be read from the flash cache or written to disk. To flush...
the dirty data from the flash disk to grid disks use the `ALTER GRIDDISK ... FLUSH` command.

- The `ALTER FLASHCACHE ... FLUSH` command stops new data from being written to the flash cache and then synchronizes all data in the flash cache with the hard disks. As a result, all data is removed from the flash cache. When the flash cache is re-enabled, the flash cache activity metrics are reset.
- The `CANCEL FLUSH` option terminates an earlier flush operation, and reinstates flash caching.
- The `NOWAIT` option allows the `ALTER` command to complete while the flush operation is in progress.
- By default, 5 percent of space on Extreme Flash storage servers is used for write-back flash cache.

**Example 8-37  Flushing Dirty Blocks from Flash Cell Disks**

This example shows how to flush dirty blocks from all flash cell disks.

```
CellCLI> ALTER FLASHCACHE ALL FLUSH
Flash cache on FD_00_scac01cel07 successfully altered
Flash cache on FD_01_scac01cel07 successfully altered
Flash cache on FD_02_scac01cel07 successfully altered
...
Flash cache on FD_14_scac01cel07 successfully altered
Flash cache on FD_15_scac01cel07 successfully altered
```

### 8.7.1.5 ALTER GRIDDISK

**Purpose**

The `ALTER GRIDDISK` command changes the attributes of all grid disks or specified grid disks.

**Caution:**

Before changing the name of a grid disk that belongs to an Oracle ASM disk group, ensure that the Oracle ASM disk group is offline.

**Syntax**

```
ALTER GRIDDISK { ALL [ FLASHDISK | HARDDISK ] | gdisk_name1 [, gdisk_name2] ... } 
{ attribute_filters { ACTIVE | INACTIVE | { FLUSH [NOWAIT] | CANCEL FLUSH } } | 
attribute_name = attribute_value [, attribute_name = attribute_value] ... 
[ attribute_filters ] [NOWAIT] }
```
Command Options

• The **FLASHDISK** option limits the **ALTER GRIDDISK** command to grid disks that are flash disks.

• The **HARDDISK** option limits the **ALTER GRIDDISK** command to grid disks that are hard disks.

• The **ACTIVE** option notifies CELLSRV to accept I/O as normal for the specified grid disks. The grid disks are visible to the database clients.

• The **INACTIVE** option makes the grid disks visible to the cell administrator, but not visible to the database clients. CELLSRV treats the grid disks as if they were offline. This mode allows management operations on the grid disks. You can do upgrading and testing on the grid disks before making the grid disks visible to database users. This functionality is similar to starting up a database in **RESTRICTED** mode.

  **Note:**

  When a grid disk that is currently in use by a database client is made **INACTIVE**, Oracle ASM takes the corresponding Oracle ASM disk offline when I/Os to the disk fail. To make the disk usable again, make the grid disk **ACTIVE** in the cell, and then bring the corresponding Oracle ASM disk back online in Oracle ASM.

• The **FLUSH** option synchronizes dirty data from the flash cache to the grid disks. Dirty data is data that has not been synchronized with the grid disk. Synchronization of dirty data can be a lengthy process, depending on the number of bytes to be synchronized. Use the following command to check the progress:

  ```
  LIST GRIDDISK ATTRIBUTES name, flushstatus
  ```

• The **CANCEL FLUSH** option terminates an earlier d flush operation, and reinstates caching.

• The **NOWAIT** option allows the **ALTER** command to complete while a resize or flush operation continues, for example:

  ```
  ALTER GRIDDISK size=5T WHERE name like 'RECO.*'
  ```

  You can use the **WHERE** clause with either a list of grid disks or with the ALL (**FLASHDISK** | **HARDDISK**) option. If you use the **WHERE** clause and do not specify
either a list of grid disks or the ALL {FLASHDISK|HARDDISK} option, then the WHERE clause acts against all disks. Some examples of this syntax are:

```
ALTER GRIDDISK WHERE name like 'DATA.*' INACTIVE

ALTER GRIDDISK ALL FLASHDISK FLUSH NOWAIT
```

**Usage Notes**

The attributes that can be changed with the `ALTER GRIDDISK` command are shown as **modifiable** in Example 8-99.

- The length of a grid disk name is limited to 30 characters.
- The `FLUSH` option stops new data from being cached on the specified grid disks until CELLSRV restarts, or the flush operation is canceled.
- The `FLUSH` option is valid for write back disks, not write through disks.
- The `size` attribute can be specified to expand or reduce space allocated to a grid disk. The corresponding Oracle ASM disk must be resized separately. The `size` attribute is specified as a number in bytes, unless the suffix M (megabytes) or G (gigabytes) is included with the number value. Grid disk space is allocated in 16 MB units, referred to as allocation units. The actual size allocated is the size of the largest multiple of allocation units less than or equal to the specified size. The minimum value is 16 MB. Values less than 16 MB are rounded up to 16 MB.
- A grid disk should not be renamed when the grid disk is being accessed. If you try to rename a grid disk when it is being accessed, then the operation fails. You can make the grid disk inactive or dismount the Oracle ASM disk group to stop access to the grid disk before renaming it.
- When an interleaved grid disk is resized, the contents of the grid disk are moved to achieve the interleaved space allocation across the cell disk. The resizing operation can take a few minutes. You can choose to have the data movement proceed as a background process by using the `NOWAIT` option. Use the `LIST GRIDDISK` command to check the status.

*Note:*

Interleaved grid disks are deprecated in Oracle Exadata System Software release 19.1.0.

- The `cachingPolicy` attribute is used to change the flash caching policy of a grid disk. Before changing the caching policy from `default` to `none`, ensure there is no cached data in flash cache for the grid disk by using the `ALTER GRIDDISK ... FLUSH` command.

Flash cache is not used with data files placed in a disk group composed of grid disks with their `cachingPolicy` set to `none`.

OEDA configures the RECO disk group to have its grid disk's `cachingPolicy` set to `none`; therefore flash cache is not used for any data files placed in RECO.

Write back caching can be disabled on the grid disks that do not need caching, such as the grid disks in the RECO disk group. This allows other objects to use...
the cache space. Use the following commands to flush the cache, and then set the cachingPolicy attribute to none:

```
ALTER GRIDDISK grid_disk_name FLUSH
ALTER GRIDDISK grid_disk_name CACHINGPOLICY="none"
```

To re-enable caching on the grid disk, do the following:

```
ALTER GRIDDISK grid_disk_name CANCEL FLUSH
ALTER GRIDDISK grid_disk_name CACHINGPOLICY="default"
```

*grid_disk_name* is the name of the grid disk.

**Example 8-38  Altering Grid Disk Attributes**

This example shows the ALTER command with the GRIDDISK object.

```
CellCLI> ALTER GRIDDISK data1_CD_01_cell01, data2_CD_01_cell101
       comment = "This grid disk is on cell01"
```

```
CellCLI> ALTER GRIDDISK ALL INACTIVE
```

**Related Topics**

- **Restrictions on Values of Common Attributes**
  Review the following restrictions for the values of attributes used by multiple CellCLI objects.

- **Attribute Filters in LIST and ALTER GRIDDISK Commands**
  You can specify which objects to display for the LIST and ALTER GRIDDISK commands with the optional *attribute_filters* clause.

### 8.7.1.6 ALTER IBPORT

**Purpose**

The ALTER IBPORT command performs an action on all InfiniBand Network Fabric ports, or specified InfiniBand Network Fabric ports.

**Note:**

This command does not apply to Oracle Exadata Database Machine X8M systems.

**Syntax**

```
ALTER IBPORT {ALL | ibport_name [, ibport_name] ...} RESET COUNTERS
```

**Usage Notes**

The *RESET COUNTERS* option resets all counters on the InfiniBand Network Fabric port.
Example 8-39  Altering IBPORT Attributes

This example shows the ALTER command with the IBPORT object.

CellCLI> ALTER IBPORT ALL RESET COUNTERS
InfiniBand Port HCA-1:1 successfully altered.
InfiniBand Port HCA-1:2 successfully altered.

CellCLI> ALTER IBPORT "HCA-1:1" RESET COUNTERS
InfiniBand Port HCA-1:1 successfully altered.

8.7.1.7 ALTER IORMPLAN

The ALTER IORMPLAN command updates the interdatabase plan and category plans for the cell.

Purpose

The directives determine the allocation of I/O resources to database clients. The ALTER IORMPLAN command is used to control disk and flash I/Os by database clients.

Syntax

ALTER IORMPLAN [dbPlan=(directive1[, directive2]...),]
[catPlan=( directive1[, directive2]...)]
[objective='iorm_objective']

- Parameters for ALTER IORMPLAN Command
- Directives for dbPlan
- Directives for catPlan

allocation and level Directives
Use the level and allocation directives to specify percentages for I/O distribution for each database.

- asmcluster Directive for dbPlan
- flashcache Directive for dbPlan
- flashcachelimit Directive for dbPlan
- flashcachemin Directive for dbPlan
- flashcachesize Directive for dbPlan
- flashlog Directive for dbPlan
  The flashLog directive is used to prevent a database from using the Flash Log.

- limit Directive
  Use the limit directive to restrict the I/O utilization for a database to a specified utilization limit.

- name Directive
  Use the name directive to specify the name of the database or category.

- pmemcache Directive for dbPlan
8.7.1.7.1 Parameters for ALTER IORMPLAN Command

Parameters

- **dbplan:** Specifies the interdatabase plan, allowing you to manage resource allocations among databases. You can activate the interdatabase plan by setting the `dbplan` attribute. If the `dbplan` attribute is not set, then the default is that each database has an equal percentage or share of the resources.

- **catplan:** Specifies the category plan, allowing you to allocate resources primarily by the category of the work being done. If the `catPlan` directive is not set, then the default is that no category plan is used. The resources are divided among databases.

- **objective:** The `objective` attribute specifies the optimization mode for IORM. By default, the objective is set to `basic` because IORM is always minimally enabled. The `objective` attribute can have the following values:
  - **auto:** Use this setting to have IORM determine the best mode based on active workloads and resource plans. IORM continuously and dynamically determines the optimization objective, based on the workloads observed, and resource plans enabled. This is the recommended value that users should start with.
  - **high_throughput:** Use this setting for critical DSS workloads that require high throughput. This setting improves throughput at the cost of I/O latency.
  - **low_latency:** Use this setting for critical OLTP workloads that require extremely good disk latency. This setting provides the lowest possible latency at the cost of throughput by significantly limiting disk utilization.
  - **balanced:** Use this setting for critical OLTP and DSS workloads. This setting balances low disk latency and high throughput. This setting limits disk utilization of large I/Os to a lesser extent than `low_latency` to achieve a balance between latency and throughput.
  - **basic:** Use this setting to disable I/O prioritization and limit the maximum small I/O latency.

To enable user-defined plans, the `objective` attribute must be set to a value other than `basic`.

Usage Notes

- **To reset a dbPlan or catPlan parameter to the default value,** set the attribute to an empty string. Single quotation marks or double quotation marks can be used to set an empty string. The marks must match, for example "" is correct, but "" is incorrect.
• The ALTER IORMPLAN command must specify at least one of the following directives: allocation, share, limit, flashcache, flashlog, flashCachelimit, flashCacheSize, or flashCacheMin.

• Consider running ALTER IORMPLAN commands as scripts because of the length of the commands.

Example 8-40  Altering IORMPLAN Objective

This example shows the ALTER command with the IORMPLAN object.

CellCLI> ALTER IORMPLAN objective='low_latency'

CellCLI> ALTER IORMPLAN objective='auto'

Example 8-41  Resetting IORMPLAN Attributes

This example shows how to reset the IORMPLAN attributes for both the dbPlan and catPlan, or individually.

CellCLI> ALTER IORMPLAN dbPlan="", catPlan=""

CellCLI> ALTER IORMPLAN dbPlan=""

CellCLI> ALTER IORMPLAN catPlan=""

8.7.1.7.2 Directives for dbPlan

The directives for a database plan (dbPlan) use the following syntax:

{(name=db_name, [share=number] [,level=number, allocation=number] [,limit=number] [, flashcache={on|off}] [, flashLog={on|off}] [, flashcachemin=number] [, flashcachelimit=number] [, pmemcachelimit=number] [, pmemcache={on|off}] [, asmcluster=asm_cluster_name] [, role={primary|standby}])

A database plan can be defined using one of the following configurations:

• share: Use the share directive to specify the relative priority for a database. A higher share value implies higher priority and more guaranteed I/O resources.

• allocation: Use the allocation attribute to specify percentages for I/O distribution for each database. You can allocate resources across different levels, using the level attribute.

• limit: Use the limit attribute to restrict the I/O utilization for a database to a specified utilization limit. The limit attribute specifies the maximum disk I/O utilization limit for databases.

To enable disk I/O management, specify share, limit, or level and allocation.
Usage Notes

Up to 1024 directives are allowed for each `dbPlan`. These directives include the mandatory `OTHER` database directive.

Note:

Only 32 directives are allowed when using the `level` and `allocation` attributes.

The following items should be checked before implementing an interdatabase plan:

- A `dbPlan` must have at least one directive for `other` when using the `allocation` attribute.
- You cannot specify values for both `allocation` and `share` directive for the same database.
- Only a single assignment of an attribute is allowed for each database and each profile.

8.7.1.7.3 Directives for catPlan

The directives for a category plan (`catPlan`) use the following syntax:

```
(name=category_name, [,level=number, allocation=number])
```

A category plan can be defined using the following configuration:

- `allocation`: Use the `allocation` directive to specify percentages for I/O distribution for each category. The `allocation` directive allows IORM to manage up to 32 categories. You can allocate resources across different levels, using the `level` directive. `level` can have a value from 1 to 8. For each `level` the total `allocation` value cannot exceed 100.

Usage Notes

You can use `allocations` to specify percentages for I/O distribution for each category. With allocations, IORM can manage up to 32 categories. You can allocate resources across 8 different `levels`. For example, you can put batch operations on a different level from sales operations.

The following items should be checked before implementing a category plan:

- You can use up to 32 directives for a `catPlan`.
- Only a single assignment of an attribute is allowed for each category and each profile.
- A `catPlan` must have at least one directive for `other set`.
- A profile name cannot be `OTHER` or `DEFAULT`. Those names are reserved words.
- `pmemcachesize`, `pmemcachemin`, `pmemcachelimit`, `flashcachesize`, `flashcachemin` and `flashcachelimit` cannot be used in `catPlan`. 
• The `flashCache` and `pmemCache` directives cannot be used with `catPlan`.
• The `flashLog` and `pmemLog` directives cannot be used with `catPlan`.
• `asmcluster` is not allowed in `catPlan` directives.
• The `role` and `type` directives are not permitted with `catPlan`.
• The `catPlan` parameter cannot be set if the `dbPlan` parameter includes the `type=profile` directive.

8.7.1.7.4 allocation and level Directives

Use the `level` and `allocation` directives to specify percentages for I/O distribution for each database.

**Purpose**

Use the `allocation` directive to specify percentages for I/O distribution for each database. You can allocate resources across different levels, using the `level` attribute.

**Syntax**

```
ALTER IORMPLAN
dbplan=(name=db_name, level=number, allocation=number,
[, directive=value, ...])
```

**Usage Notes**

• You must specify values for both `level` and `allocation`.
• Use the `level` directive to allocate resources across multiple levels, with a maximum of 8 levels.
• You cannot specify a value greater than 100 for total `allocation` at each level.
• You cannot have more than 32 directives for `dbPlan` when the `allocation` directive is used.
• You must specify at least one directive with `name=OTHER` for `dbPlan` when using the `allocation` directive.
• You cannot specify the `allocation` directive with `name=default`.
• You cannot specify `asmcluster` if also specifying `allocation`.

**Example 8-42  Altering IORMPLAN Attributes**

These examples show the `ALTER` command with the `allocation` directive.

```
CellCLI> ALTER IORMPLAN
    -
catPlan=(
    - (name=administrative, level=1, allocation=80),
    - (name=interactive, level=2, allocation=90),
    - (name=batch, level=3, allocation=80),
    - (name=maintenance, level=4, allocation=50),
    - (name=other, level=4, allocation=50)
    - ),
    -
dbPlan=((name=sales_prod, level=1, allocation=80),
    - (name=finance_prod, level=1, allocation=20),
    -
```

8.7.1.7.5 asmcluster Directive for dbPlan

Purpose
Starting with Oracle Exadata System Software release 19.1.0, you can use the asmcluster directive to specify different directives for databases with the same unique name running in different Oracle ASM clusters.

Syntax
ALTER IORMPLAN dbplan=(name=name[,...],
amascluster=asm_cluster_name[,...])

Usage Notes
- The value of the asmcluster directive should match the asm field in the cellkey.ora file when setting up ASM-scoped security.
- The asmcluster directive cannot be used with DEFAULT and OTHER.
- The asmcluster directive cannot be used with catPlan.
- You cannot use asmcluster with allocation and level.

Example 8-43 Specifying an Oracle ASM Cluster Using IORMPLAN Directives
For a multi-cluster environment with ASM-scoped security mode, specify different directives for databases with the same unique name using the asmcluster directive.

CellCLI> ALTER
  IORMPLAN
    dbPlan=((name=pdb1, share=4, flashCacheMin=5G),
       asmascluster=cluster1),-
        (name=pdb1, share=2, limit=80, asmascluster=cluster2), -
        (name=pdb2, share=2, flashcachelimit=2G,
       asmascluster=cluster1),-
        (name=default, share=1, flashcachelimit=1G))

8.7.1.7.6 flashcache Directive for dbPlan

Purpose
The flashcache directive is used to prevent a database from using the Flash Cache. This ensures that Flash Cache space is reserved for mission-critical databases.
Note:

If the `flashcache` attribute is not specifically set to `off` for a database, then that database uses Flash Cache by default.

Syntax

ALTER IORMPLAN dbPlan=
( name=db_name, ...[, flashcache={on|off}] [, ...])

Usage Notes

- The `flashcache` directive cannot be used with `catPlan`.
- You can specify `flashcache` directives for a profile.
- You cannot use the `flashcache` directive with `OTHER`.
- The `flashcache` directive cannot be set to `OFF` when specifying the `flashcachemin`, `flashcachelimit`, or `flashcachesize` directives.

Example 8-44 Setting Flash Cache Use with IORMPLAN Directives

These examples show how to enable Flash Cache use with the IORMPLAN directives.

CellCLI> ALTER IORMPLAN
  dbPlan=((name=sales_prod, flashCache=on),
         (name=sales_dev, flashCache=on),
         (name=sales_test, flashCache=off),
         (name=other, flashCache=off))

8.7.1.7.7 flashcachelimit Directive for dbPlan

Purpose

The `flashcachelimit` directive is a soft limit that specifies the maximum space in the Flash Cache that a database can use. If the Flash Cache is not full, the database can exceed its allocated quota.

Syntax

ALTER IORMPLAN dbPlan=
( name=db_name[, ...], flashcachelimit=number [, ...])

Usage Notes

- The `flashcachelimit` directive cannot be set to 0.
- The value for `flashcachelimit` must be at least 4 MB.
- You specify the value for `flashcachelimit` in bytes. You can also use the suffix `M` (megabytes), `G` (gigabytes), or `T` (terabytes) to specify larger values, for example, `300M`, `150G`, or `1T`. 
• The `flashcachesize` directive cannot be specified for a database if `flashcachelimit` is specified.

• The sum of `flashcachesize` and `flashcachelimit` can't be less than `flashcachemin`.

• The `flashcachelimit` directive cannot be used with `catPlan`.

• The `flashcachelimit` directive cannot be used with `OTHER`.

• You can specify `flashcachelimit` directives for a profile.

### Example 8-45  Specifying Flash Cache Quotas Using IORMPLAN Directives

This example shows how to configure flash cache quotas using `IORMPLAN` directives.

```
CellCLI> ALTER IORMPLAN
             dbPlan=((name=prod, share=8,
                      flashCacheMin=4096),
                   (name=dev, share=2, flashCacheMin=1G,
                      flashCacheLimit=4G),
                   (name=test, share=1, limit=40, flashCacheLimit=2048))
```

#### 8.7.1.7.8 flashcachemin Directive for dbPlan

**Purpose**

The `flashcachemin` directive specifies the minimum guaranteed space in the Flash Cache for a database.

**Syntax**

```
ALTER IORMPLAN dbPlan=
              ( name=db_name[, ...] , flashcachemin=number [, ...])
```

**Usage Notes**

• You specify the value for `flashcachemin` in bytes. You can also use the suffix M (megabytes), G (gigabytes), or T (terabytes) to specify larger values, for example, 300M, 150G, or 1T.

• The `flashcachemin` directive cannot be set to 0.

• The value for `flashcachemin` must be at least 4 MB.

• The sum of `flashcachemin` across all the directives for a plan should be less than the size of Flash Cache.

• The sum of `flashcachesize` and `flashcachelimit` across all the directives in a plan can't be less than `flashcachemin`.

• The `flashcachemin` directive cannot be used with `OTHER`.

• The `flashcachemin` directive cannot be used with `catPlan`.

• You can specify `flashcachemin` directives for a profile.
8.7.1.7.9 flashcachesize Directive for dbPlan

Purpose
Starting with Oracle Exadata System Software release 12.1.2.2.0, the `flashcachesize` directive specifies the guaranteed hard limit space in the flash cache for a database. This space in the flash cache is reserved but since this is a hard limit the database cannot exceed its quota even when the flash cache is not full.

Beginning with Oracle Exadata System Software release 19.1, if both `flashcachemin` and `flashcachesize` are not set for the same database in an interdatabase plan, and the sum of all `flashcachemin` and `flashcacheSize` directives is bigger than the current flash cache size, then the system is considered to be over-provisioned. If the system is over-provisioned, then the `flashcachemin` and `flashcachesize` values can not be guaranteed.

Syntax

```
ALTER IORMPLAN dbPlan=
  ( name=db_name, [ , ...] flashcachesize=number [ , ...])
```

Usage Notes

- You specify the value for `flashcachesize` in bytes. You can also use the suffix M (megabytes), G (gigabytes), or T (terabytes) to specify larger values, for example, 300M, 150G, or 1T.
- The `flashcachesize` directive cannot be set to 0.
- The value for `flashcachesize` must be at least 4 MB.
- The `flashcachesize` directive cannot be specified for a database if `flashcachelimit` is specified.
- The sum of `flashcachesize` and `flashcachelimit` across all the directives in a plan can't be less than `flashcachemin`.
- The `flashcachesize` directive cannot be used with `OTHER`.
- The `flashcachesize` directive cannot be used with `catPlan`.
- You can specify `flashcachesize` directives for a profile.

8.7.1.7.10 flashlog Directive for dbPlan

The `flashlog` directive is used to prevent a database from using the Flash Log.

Syntax

```
ALTER IORMPLAN dbPlan=
  ( name=db_name, ...[, flashlog={on|off}] [ , ...])
```

Usage Notes

- The `flashlog` directive cannot be used with `catPlan`.
- You can specify the `flashlog` directive for a profile.
• You cannot use the `flashlog` directive with `OTHER`.

**Note:**
If the `flashLog` directive is not specifically set to `off` for a database, then that database uses Flash Log by default.

**Example 8-46  Setting Flash Log Use with IORMPLAN Directives**
These examples show how to enable Flash Log use with the `IORMPLAN` directives.

```
CellCLI> ALTER IORMPLAN
     dbPlan=(
        (name=oltp, level=1, allocation=80, flashCache=on, flashLog=on),
        (name=dss, level=1, allocation=20, limit=50, flashCache=off, flashLog=off),
        (name=other, level=2, allocation=100, flashCache=off, flashLog=off))
```

---

**8.7.1.7.11 limit Directive**

Use the `limit` directive to restrict the I/O utilization for a database to a specified utilization limit.

**Purpose**

Use the `limit` directive to restrict the I/O utilization for a database to a specified utilization limit. The `limit` directive specifies the maximum disk I/O utilization limit for databases. This resource allocation method is ideal for pay-for-performance use cases and should not be used to implement fairness.

**Syntax**

```
ALTER IORMPLAN
dbplan=( name=db_name, limit=number [, directive=value, ...])
```

**Usage Notes**

• Use a value from 1 to 100 for `limit` to specify the maximum disk I/O utilization limit for databases. If a maximum utilization limit is specified, then excess capacity is never used by the databases. As a result, it is possible that the disks are running below full capacity when maximum utilization limits are specified.

**Note:**
Specifying low `limit` values can have a significant performance impact and is not advisable for regular consolidation use cases.
• When using Resource Manager to manage resources for a single database, the `MAX_UTILIZATION_LIMIT` attribute specifies the disk I/O utilization limit for consumer groups. Refer to *Oracle Database Administrator’s Guide* and *Oracle Database PL/SQL Packages and Types Reference* for additional information about this attribute.

• You can use `name=OTHER` with the limit directive. You can also use `name=default` with the limit directive.

• The limit directive cannot be used with `catPlan`.

**Example 8-47 Using the name Directive with `dbPlan`**

```
CellCLI> ALTER IORMPLAN                            -
    dbplan=((name=db1, limit=40),              -
        (name=db2, limit=40),              -
        (name=other, limit=20))
```

8.7.1.7.12 name Directive

Use the name directive to specify the name of the database or category.

**Syntax**

```
ALTER IORMPLAN
    dbPlan=( name=db_name, [, directive=value, ...])
```

```
ALTER IORMPLAN
    catPlan=( name=category_name, [, directive=value, ...])
```

**Usage Notes**

The database or category name cannot start with an underscore (_). Each database or category name must be followed by at least one directive, for example:

• `(name=sales, share=8)`
• `(name=oltpdg, limit=80)`
• `(name=dwh, flashcachesize=50G)`

There are two special values that can be used with `name`:

• The `OTHER` value is a distinguished name. It refers to all clients that are not specifically named in other directives. Both `catPlan` and `dbPlan` must include at least one directive for `OTHER`.

• The `DEFAULT` value is a distinguished name. It provides default values for all databases or categories not explicitly named in a plan so you do not have to add directives for each client. Any database that does not have an explicit directive uses the values that are specified for the `DEFAULT` database name.

**Usage Notes**

• For `catPlan`, you can use the `OTHER` directive only with level and allocation directives.
For dbPlan, you can use the OTHER directive with level and allocation directives, or with the limit directive.

You must specify at least one directive with name=other for dbPlan when using the allocation directive or the CellCLI utility returns an error.

You cannot use role, asmcluster, flashcache, flashlog, flashcachemin, flashcachelimt, flashcachesize, pmemcache, pmemlog, pmemcachemin, pmemcachelimt, or pmemcachesize directives with OTHER.

You cannot use the share directive with OTHER.

You cannot specify level, allocation, asmcluster or role directives for DEFAULT.

Example 8-48 Using the name Directive with dbPlan

CellCLI> ALTER IORMPLAN
      dbplan=((name=db1, limit=50),
      (name=db2, limit=50),
      (name=other, level=1, allocation=25))

Example 8-49 Setting a Database Plan using the default Directive

This example shows how to use the default directive to set the share allocation for all databases except dev01 and dev02.

ALTER IORMPLAN
      (name=dev01, share=1, limit=50, flashLog=off),
      (name=dev02, share=1, limit=25, flashCache=off)
      (name=default, share=4))

8.7.1.7.13 pmemcache Directive for dbPlan

Purpose

The pmemcache directive is used to prevent a database from using the PMEM Cache. This ensures that PMEM Cache space is reserved for mission-critical databases.

Note:

If the pmemcache directive is not specifically set to off for a database, then that database uses PMEM Cache by default.

Syntax

ALTER IORMPLAN dbPlan=
      ( name=db_name, ...[, pmemcache={on|off}] [, ...])

Usage Notes

• The pmemcache directive cannot be used with catPlan.

• You can specify pmemcache directives for a profile.

• You cannot use the pmemcache directive with OTHER.
The `pmemcache` directive cannot be set to `OFF` when specifying the `pmemcachemin`, `pmemcachelimit`, or `pmemcachesize` directives.

The `pmemcache` directive is not set to `off` if the `pmemcachemin`, `pmemcachelimit`, or `pmemcachesize` directives are specified.

Example 8-50  Setting PMEM Cache Use with IORMPLAN Directives

These examples show how to enable PMEM Cache use with the IORMPLAN directives.

CellCLI> ALTER IORMPLAN
         dbPlan=((name=sales_prod, pmemCache=on),
         (name=sales_dev, pmemCache=on),
         (name=sales_test, pmemCache=off),
         (name=other, pmemCache=off))

8.7.1.7.14 pmemcachelimit Directive for dbPlan

Purpose

The `pmemcachelimit` directive is a soft limit that specifies the maximum space in the PMEM Cache that a database can use. If the PMEM Cache is not full, the database can exceed its allocated quota.

Syntax

```
ALTER IORMPLAN dbPlan=
    { name=db_name[ , ...], pmemcachelimit=number [, ...] }
```

Usage Notes

- The `pmemcachelimit` directive cannot be set to 0.
- The value for `pmemcachelimit` must be at least 4 MB.
- You specify the value for `pmemcachelimit` in bytes. You can also use the suffix M (megabytes), G (gigabytes), or T (terabytes) to specify larger values, for example, 300M, 150G, or 1T.
- The `pmemcachesize` directive cannot be specified for a database if `pmemcachelimit` is specified.
- The sum of `pmemcachesize` and `pmemcachelimit` can't be less than `pmemcachemin`.
- The `pmemcachelimit` directive cannot be used with `catPlan`.
- The `pmemcachelimit` directive cannot be used with `OTHER`.
- You can specify `pmemcachelimit` directives for a profile.

8.7.1.7.15 pmemcachemin Directive for dbPlan

Purpose

The `pmemcachemin` directive specifies the minimum guaranteed space in the PMEM Cache for a database.
### 8.7.1.7.16 pmemcachesize Directive for dbPlan

**Purpose**

Starting with Oracle Exadata System Software release 19.3.0, the `pmemcachesize` directive specifies a limit on space usage in the PMEM cache for a database. If both `pmemcachemin` and `pmemcachesize` are not set for the same database in an interdatabase plan, and the sum of all `pmemcachemin` and `pmemcachesize` directives is bigger than current PMEM cache size, then the system is considered to be over-provisioned. The `pmemcachemin` and `pmemcachesize` values cannot be guaranteed in an over-provisioned system.

**Syntax**

```
ALTER IORMPLAN dbPlan=
( name=db_name[, ...] , pmemcachesize=number [, ...])
```

**Usage Notes**

- You specify the value for `pmemcachesize` in bytes. You can also use the suffix `M` (megabytes), `G` (gigabytes), or `T` (terabytes) to specify larger values, for example, `300M`, `150G`, or `1T`.
- The `pmemcachesize` directive cannot be set to 0.
- The value for `pmemcachesize` must be at least 4 MB.
- The sum of `pmemcachemin` across all the directives for a plan should be less than the size of PMEM Cache.
- The sum of `pmemcachesize` and `pmemcachelimit` across all the directives in a plan can't be less than `pmemcachemin`.
- The `pmemcachesize` directive cannot be used with `OTHER`.
- The `pmemcachesize` directive cannot be used with `catPlan`.
- You can specify `pmemcachesize` directives for a profile.
The sum of `pmemcachemin` directives for a plan can’t be greater than `pmemcachemin`.

The sum of `pmemcachemin` and `pmemcachelimit` across all the directives in a plan can’t be less than `pmemcachemin`.

The `pmemcachemin` directive cannot be used with `OTHER`.

The `pmemcachemin` directive cannot be used with `catPlan`.

You can specify `pmemcachemin` directives for a profile.

8.7.1.7.17 pmemlog Directive

**Purpose**

The `pmemlog` directive is used to prevent a database or category from using the PMEM Log. If the `pmemLog` directive is not specifically set to `off` for a database, then that database uses PMEM Log by default.

**Syntax**

```
ALTER IORMPLAN dbPlan=
(name=db_name, ...[, pmemLog={on|off}] [, ...])
```

**Usage Notes**

- The `pmemLog` directive cannot be used with `catPlan`.
- You cannot use the `pmemLog` directive with `OTHER`.
- You can specify `pmemLog` directives for a profile.

**Example 8-51  Setting PMEM Log Use with IORMPLAN Directives**

These examples show how to enable PMEM Log use with the `IORMPLAN` directives.

```
CellCLI> ALTER IORMPLAN dbPlan=
(name=oltp, level=1, allocation=80, pmemCache=on, pmemLog=on),
(name=dss, level=1, allocation=20, limit=50, pmemCache=off, pmemLog=on),
(name=other, level=2, allocation=100, pmemCache=off, pmemLog=on))
```

8.7.1.7.18 role Directive for dbPlan

**Purpose**

The `role` directive enables you specify different plan directives based on the Oracle Data Guard database role. The directives for the database are applied only when the databases are in the specified database role. In the case of a switchover or fail over, the new directive values are automatically applied by IORM.
Syntax

ALTER IORMPLAN dbplan=(name=name, ..., role={primary | standby})

Usage Notes

- The role directive must be the last directive specified in the directive.
- If the role directive is not specified, then the other directives are applied regardless of the database role.
- The role directive cannot be used with catPlan.
- When using the allocation method, for databases that have the same value for the role directive, the sum of all allocation values, including OTHER, cannot exceed 100.
- The role directive cannot be used with DEFAULT and OTHER.

Example 8-52  Using the role Directive with dbPlan Allocation Method

CellCLI> ALTER IORMPLAN -
     dbPlan=((name='sales_prod', level=1, allocation=30, role=primary), -
     (name=sales1, level=1, allocation=20, role=standby), -
     (name=sales2, level=1, allocation=20, role=standby), -
     (name=other, level=3, allocation = 50))

Example 8-53  Using the role Directive with dbPlan Share Method

ALTER IORMPLAN -
     dbPlan=((name=salesprod, share=4, role=primary), -
     (name=salesprod, share=1, limit=50, role=standby) -
     (name=finance, share=4) -
     (name=hr, share=2))

8.7.1.7.19 share Directive for dbPlan

Purpose

Use the share directive to specify the relative priority for a database. A higher share value implies higher priority and more guaranteed I/O resources. This is the recommended method for interdatabase plans.

Syntax

ALTER IORMPLAN
     dbplan=( name=db_name, share=number [, directive=value, ...])
Usage Notes

- The share value is 1 to 32, with 1 being the lowest share, and 32 being the highest share. A higher share value implies higher priority and more guaranteed I/O resources. The sum of all share directives in a plan cannot be greater than 32768.
- The share directive allows IORM to manage up to 1024 databases.
- You can use share with the type=profile directive.
- You cannot use share with OTHER.
- The share directive cannot be used with catPlan.

Example 8-54 Setting a Database Plan Using the share Directive

This example shows how to configure dbPlan using the share directive.

```
ALTER IORMPLAN
  dbplan=((name=sales01, share=4),
          (name=sales02, share=4),
          (name=sales03, share=4),
          (name=fin01, share=4),
          (name=fin02, share=4),
          (name=fin03, share=4),
          (name=dev01, share=1, limit=50, flashLog=off),
          (name=dev02, share=1, limit=25, flashCache=off))
```

8.7.1.7.20 type Directive for dbPlan

Purpose

The type directive enables you to create a profile, or template, to ease management and configuration of resource plans in large consolidation environments with hundreds of databases. You use the type=profile directive to specify that the directives are for a profile rather than a specific database. Then, for those databases that you want to use a specific profile, you set the database initialization parameter db_performance_profile to the value of name that was used for a specific profile and restart the databases.

If type is not specified, then the directive defaults to the database directive. The dbplan can contain a combination of profile and individual database directives.

Syntax

```
ALTER IORMPLAN dbplan=(name=name[,...], type={database | profile}[,...])
```

Usage Notes

- The type directive must be set to either database or profile.
- When using type=profile, the name directive is used as the identifier for the profile.
- The type directive cannot be used with catPlan.
You can specify `share`, `limit`, `flashcache`, `flashlog`, `flashcachemin`, `flashcachelim`, `flashcachesize`, `pmemcachemin`, `pmemcachelimit`, or `pmemcachesize` attributes for a profile.

You cannot use `level`, `allocation`, and `role` with `type`.

Databases that map to a particular profile inherit the directive values specified for that profile.

A profile name cannot be `OTHER` or `DEFAULT`. Those names are reserved words.

Example 8-55  Creating a Profile with the type Directive

This example shows how to specify a profile as part of a `dbPlan`.

```
ALTER IORMPLAN DBPLAN=((name=gold, share=10, limit=100, type=profile),  
(name=silver, share=5, limit=60, type=profile),  
(name=bronze, share=1, limit=20, type=profile))
```

8.7.1.8 ALTER LUN

**Purpose**

The `ALTER LUN` command re-enables all LUNs or specified LUNs.

**Syntax**

```
ALTER LUN { ALL | lun1 [ , lun2 ] ... }  
       REENABLE FORCE
```

**Usage Notes**

This command creates the cell disk and grid disk metadata on a replacement disk.

This command rebuilds redundancy for the system area of the system disks even when the system LUN is in a normal state.

⚠️ Caution:

Data might be lost when using this command.

Example 8-56  Re-enabling a LUN

This example shows the `ALTER` command with the `LUN` object.

CellCLI> ALTER LUN 'x:7' REENABLE FORCE

CellCLI> ALTER LUN ALL REENABLE FORCE
8.7.1.9 ALTER OFFLOADGROUP

Purpose

The `ALTER OFFLOADGROUP` command enables you to alter modifiable attributes of offload groups, and also to restart, start up, and shut down services.

Syntax

```
ALTER OFFLOADGROUP { offloadgroup1 [,offloadgroup2, ...] } 
{attribute_name = attribute_value [, attribute_name = attribute_value ...]} 
| STARTUP | RESTART | SHUTDOWN }
```

Usage Notes

- The `offloadgroupN` (where N is a number) parameters specify the names of the offload groups whose attributes you want to modify, or that you want to start, shut down, or restart.
- The `attribute_name` and `attribute_value` parameters specify the name and value of the attribute you want to modify.
- The `STARTUP` parameter specifies that the offload group(s) is to be started.
- The `RESTART` parameter specifies that the offload group(s) is to be shut down, then started.
- The `SHUTDOWN` parameter specifies that the offload group(s) is to be shut down.

Examples

Example 8-57  Updating the "Comment" Attribute

```
ALTER OFFLOADGROUP offloadgroup1 comment='System group'
```

Example 8-58  Starting up the Offload GroupNamed "offloadgroup1"

```
ALTER OFFLOADGROUP offloadgroup1 startup
```

Related Topics

- LIST OFFLOADGROUP

8.7.1.10 ALTER PHYSICALDISK

Purpose

The `ALTER PHYSICALDISK` command prepares a disk for replacement.

Syntax

```
ALTER PHYSICALDISK { ALL [harddisk] | disk_id1 [,disk_id2] ... } 
{ DROP FOR REPLACEMENT [FORCE] | REENABLE }
```
Usage Notes

The DROP FOR REPLACEMENT option:

- Is supported only for hot-pluggable disks
- Checks if it is safe to proactively replace the specified disk. For example, if you attempt to drop the last good system disk, then replacing it would cause the system to crash.
- Offlines the data grid disks on the physical disk if they exist
- Flushes the flash cache store on the disk if it's part of flash cache
- Flushes the disk controller cache if the disk is a hard disk
- Prepares the device so that it can be removed online. For example, for flash devices this option powers off the associated PCIe slot.
- If you use the FORCE option with DROP FOR REPLACEMENT, then this allows the command to proceed even if the software deems it unsafe.

The REENABLE option re-enables a normal physical disk that was dropped for replacement.

The SERVICELED option is now obsolete. If you use this option, you will get the error message CELL-04591.

Examples

Example 8-59  Dropping a Normal, Functioning Physical Disk

This example shows how to drop a physical disk.

CellCLI> ALTER PHYSICALDISK FLASH_5_1 DROP FOR REPLACEMENT

Example 8-60  Re-enabling a Physical Disk

This example shows how to re-enable a physical disk.

CellCLI> ALTER PHYSICALDISK 12:3 REENABLE

Related Topics

- LIST PHYSICALDISK
- About Quoting Object Names
  If an object name has characters that cause parsing errors enclose the object name in quotes.

8.7.1.11 ALTER PMEMCACHE

Purpose

The ALTER PMEMCACHE command can alter the set of cell disks used by PMEM cache, flush dirty blocks from PMEM cache, or cancel a previous flush operation on the specified cell disks to re-enable caching.
Syntax

ALTER PMEMCACHE { ALL | CELLDISK="cdisk1 [,cdisk2] ..." [FORCE]}  
{FLUSH [NOWAIT] | CANCEL FLUSH}

Usage Notes

- The ALL option affects all available PMEM cell disks.
- The CELLDISK option allows you to specify individual cell disks. cdiskn represents a cell disk name.
- The FLUSH option synchronizes dirty data from the PMEM cache to the cell disks. Dirty data is data that has been modified in the cache but not yet synchronized with the data on disk. Synchronization of dirty data can be a lengthy process, depending on the number of bytes to be synchronized. Use the following command to check the progress of the flush operation:

  LIST CELLDISK ATTRIBUTES name, flushstatus, flusherror

- The ALTER PMEMCACHE ... FLUSH command stops new data from being written to the PMEM cache and then synchronizes all data in the PMEM cache with the flash disks. As a result, all data is removed from the PMEM cache.
- The CANCEL FLUSH option terminates an earlier flush operation, and reinstates PMEM caching.
- The NOWAIT option allows the ALTER command to complete while the flush operation is in progress.
- The FORCE option can be used to forcefully change the set of cell disks used by the PMEM cache.

Example 8-61 Using the ALTER PMEMCACHE Command

The following command specifies that the PMEM cache uses all PMEM cell disks.

CellCLI> ALTER PMEMCACHE ALL

The following command specifies that the PMEM cache uses only two PMEM cell disks, ignoring any errors or warnings.

CellCLI> ALTER PMEMCACHE CELLDISK='PM_01_mycell, PM_03_mycell' FORCE
The following command specifies a flush operation for all PMEM cell disks, and returns the prompt before the operation completes.

CellCLI> ALTER PMEMCACHE ALL FLUSH NOWAIT

8.7.1.12 ALTER QUARANTINE

Purpose

The ALTER QUARANTINE command changes the attributes for a quarantine.

Syntax

ALTER QUARANTINE { ALL | quarantine1 [,quarantine2] ... }
    attribute_name = attribute_value
    [, attribute_name = attribute_value]...

Usage Notes

Only modifiable fields can be changed.

Examples

The following example shows the ALTER command with the QUARANTINE object.

Example 8-62  Altering a Quarantine

CELLCLI> ALTER QUARANTINE 12 comment='bugX'

8.7.1.13 ALTER SOFTWAREUPDATE

Purpose

You can schedule software updates by setting SOFTWAREUPDATE attributes. The ALTER SOFTWAREUPDATE command enables you to alter modifiable Software Update attributes, to validate the pre-requirements for the software update, or to start the upgrade immediately.

You can also run the ALTER SOFTWAREUPDATE command using exacli.

Syntax

ALTER SOFTWAREUPDATE {VALIDATE PREREQ | UPGRADE [FORCE] | attribute = attribute value [,attribute = attribute value...]}

Options for the ALTER SOFTWAREUPDATE Command

- VALIDATE PREREQ
  
  Run software update check pre-requirement steps now. This will download the software update pre-requirement code for the update specified by the store attribute. These checks are run automatically as part of update. Use this option only if you want to run prerequisite checks explicitly. Any error found will be displayed. A stateful alert will be raised if any error is found by the VALIDATE PREREQ command.

- UPGRADE [FORCE]
Run the software update (including the pre-requisite steps) now, using the software location specified by the Software Update store attribute. Use this command if you want to perform the update now rather than wait for the time specified by the Software Update time attribute.

If FORCE is specified, then the upgrade continues despite any pre-requisite check errors.

- **attribute = attribute value**

Modify the specified Software Update attributes to the values provided.

### Attributes for SOFTWAREUPDATE

The following attributes for the ALTER SOFTWAREUPDATE command are modifiable:

- **frequency**: Storage server updates can automatically be done periodically by setting the frequency attribute to the desired frequency. You can specify one of the following values: {none | daily | weekly | biweekly }. If the value specified for frequency is "" or none, then the scheduled update is only done once. The value none can be used for the frequency in Oracle Exadata System Software release 19.1.0 or later.

- **name**: The name of the patch to use in the update, which includes the software version string such as 18.1.1.0.0.171018. If you use the ALTER SOFTWAREUPDATE UPGRADE or ALTER SOFTWAREUPDATE VALIDATE PREREQ command, then the software store is checked and the name attribute is set automatically (if not already set) to the latest available version in the software store. Otherwise, the name attribute has a value of unknown. If there are multiple software versions at the store site, then this attribute can be used to specify which version should be used.

  Patches downloaded from My Oracle Support use a different name format, for example p26875767_181100_Linux-x86-64.zip. If you are using Oracle Exadata System Software 18c (18.1.0) or 18c (18.1.1), then you must rename the downloaded patch file so it has a name like 18.1.1.0.0.171018. Starting with Oracle Exadata System Software release 18.1.2, the ALTER SOFTWAREUPDATE command accepts patch names of the form p26875767_181100.

- **store**: A URL for the location of the software update file. MS finds and downloads the software update file using the protocol specified by the URL. You can use the following protocols for the store URL: HTTP or HTTPS. The update software can be staged anywhere that is accessible using the management or RDMA Network Fabric network of the storage server.

- **time**: A future date and time at which the software update should be performed. The time can be specified as a local informal date and time, for example "1 AM, next Tuesday". If the date and time is valid then the output from setting this attribute shows the interpreted time in standard format with timezone offset, such as 2017-08-22T01:00:00-08:00.

  If you set this attribute to the empty string, "", then it cancels the scheduled software update.

- **timeLimitInMinutes**: An update may wait for other storage servers to complete their updates in order to preserve disk group redundancy. By default, there is no limit on the amount of time which can be spent waiting to update. This attribute may be set to a positive, maximum integer which represents the number of minutes a storage server will spend waiting to update. If an update does not start
within the time specified by the limit, then the update is canceled and an update alert is reported.

Usage Notes

• The store attribute directs storage servers to a URL from which they can automatically download update software. Each storage server downloads the software to its active partition, and then loads the software on its passive partition. The storage servers reboot to the new software version according to the desired schedule.

• Software download and the prerequisite check will begin up to a week before the scheduled update time.

• The update progress can be monitored by displaying the non-modifiable Software Update status attribute.

• Software updates do not occur if the upgrade software is already installed

• You can use dcli or exacli to schedule and install updates using the ALTER SOFTWAREUPDATE command.

• The Software Update feature supports using HTTPS transport for software downloads. When using HTTPS, TLS certificate checks are performed by default. If the remote server's certificate cannot be validated then the following error is reported:

CELL-00076: An error occurred during download of software update: source https://hostname:port is not available.
CELL-00092: The store's TLS certificate cannot be authenticated with known CA certificates.

This can happen if the remote server uses a self-signed certificate or if the remote server uses a certificate signed by a certificate authority (CA) that is not included in the storage server's CA store. You can use the following procedure to add a CA certificate to the storage server's CA store. This is a security setup step which requires shell access as root on the storage server.

1. Get a CA certificate that can verify the remote server. The certificate should be stored in PEM or DER file format.

2. Copy the file to the storage server at this directory: /etc/pki/ca-trust/source/anchors/

3. Run following commands:

   update-ca-trust enable

   update-ca-trust extract

Use the operating system man utility to get more information about the update-ca-trust command.
### Examples

**Example 8-63   Modifying the Software Update time Attribute**

Modify the scheduled time of the next software update to 1 a.m. on Thursday.

```
CellCLI> ALTER SOFTWAREUPDATE time = "1 AM Thursday"
CELL update is scheduled to begin at 2017-08-24T01:00:00-08:00
```

**Example 8-64   Setting the Software Update store Attribute**

This example shows how to set the store attribute to a location that uses HTTPS protocol.

```
ALTER SOFTWAREUPDATE store="https://my-exa-store/cell"
```

**Example 8-65   Starting a Software Update Immediately**

This example shows how to immediately start a software update using the attribute values already specified.

```
ALTER SOFTWAREUPDATE UPGRADE FORCE
```

### 8.7.1.14 ALTER THRESHOLD

**Purpose**

The `ALTER THRESHOLD` command updates the attribute values of all thresholds or the specified thresholds.

**Syntax**

```
ALTER THRESHOLD { ALL | threshold_name [, threshold_name ...] }
  attribute_name = attribute_value
  [, attribute_name = attribute_value]...
```

**Usage Notes**

The attributes that can be changed with the `ALTER` command are shown as modifiable in Describing the THRESHOLD Object.

**Examples**

The following example shows how to alter threshold attributes.

**Example 8-66   Altering Threshold Attributes**

```
CellCLI> ALTER THRESHOLD ct_io_wt_rq.interactive warning=10, critical=20, -
  comparison='=', occurrences=2, observation=10
```

```
CellCLI> ALTER THRESHOLD ALL occurrences=3
```

**Related Topics**

- CREATE THRESHOLD
8.7.1.15 ALTER USER

Purpose
The ALTER USER command changes the attributes of a user role.

Syntax
ALTER USER user1 attribute_name1 = attribute_value1 \ [attribute_name2 = attribute_value2, ...]

Usage Notes
- The user name cannot be root, celladmin or cellmonitor. Those are reserved.
- The user name should be unique.
- The system prompts for a password for the new user. The password must have 12 to 40 alphanumeric characters or special characters !@#$%^&*() with at least one digit, one lowercase letter, and one uppercase letter. Starting with Oracle Exadata System Software release 18.1.0.0.0, the password can be 8 to 40 characters in length and can also utilize the special characters - and _.
- The new password cannot be the same as the current password for the user.

Example 8-67 Using the ALTER USER Command
This example shows how to change a user’s password.

CellCLI> ALTER USER sjones password=TOPsecret2345

8.7.2 ASSIGN KEY

Purpose
The ASSIGN KEY command assigns or removes a security key to or from a client.

Syntax
ASSIGN KEY FOR [ASMCLUSTER] 'client_name1' = 'key-value1' [, 'client_name2' = 'key-value2'...]

ASSIGN KEY FOR CELL 'key-value'

ASSIGN KEY FOR [REMOTE | LOCAL] CELL 'client_name1' = 'key-value1' [, 'client_name2' = 'key-value2'...]

Options
- client_name is an alias that is the unique name (DB_UNIQUE_NAME) for a database client or Oracle ASM cluster.
The client name or Oracle ASM cluster name not case-sensitive. For example, ASM1 and asm1 are treated as the same value.

- **key-value** is a hexadecimal string key that is assigned to the client as a security key. The key value is generated with the `CREATE KEY` command. The key values assigned with the `ASSIGN` command must match the key in the client cellkey.ora file on the database servers. The key value can be the same for multiple clients that need the same access. An empty string for the **key-value** removes a previously assigned key.

- Starting with Oracle Exadata System Software release 12.2.1.1.0, you can use the optional keyword `ASMCLUSTER` to indicate that the client is an Oracle ASM cluster. The Oracle ASM cluster alias must not be longer than 14 characters, and only alphanumeric and hyphen characters are allowed.

- Starting with Oracle Exadata System Software release 12.2.1.1.0, the use of the `CELL` keyword can be used to assign a single key to all storage servers to enable cell-to-cell direct operations. You specify only a single **key-value**; you do not specify a `client_name`. You cannot use a list of values with the `CELL` keyword.

- Starting with Oracle Exadata System Software release 12.2.1.1.0, the `FOR LOCAL CELL` clause assigns a cell key to the local (current) cell. If you specify `FOR LOCAL CELL`, there can be only one key; a list of values is not supported. The `client_name` is a unique identifier for each cell.

- Starting with Oracle Exadata System Software release 12.2.1.1.0, the `FOR REMOTE CELL` clause specifies the cell keys that the current cell will accept. The `client_name` is the unique identifier for the cell assigned the **key-value**. You can specify a single client and key, or a list of values.

### Usage Notes

- For ASM-scoped security or DB-scoped security, the client aliases must be entered in the `availableTo` attribute of the `GRIDDISK` object.

- When using the `ASMCLUSTER` keyword in Oracle Exadata System Software release 12.2.1.1.0 or later, if you specify a client name and key that already exists (that is a key was already specified for an Oracle ASM client prior to Oracle Exadata System Software release 12.2.1.1.0), then the client will be changed to be an Oracle ASM cluster client. In this case, the name and key will be removed from the ASM-scoped security list, and added as an Oracle ASM cluster client. Grid disks with this Oracle ASM client in their ACL can remain online for this operation.

### Examples

**Example 8-68 Assigning Keys to Clients**

This example shows how to use the `ASSIGN KEY` command to assign keys to one or multiple clients.

```
CellCLI> ASSIGN KEY FOR 'db0' = 'b67d5587fe728118af47c57ab8da650a'
CellCLI> ASSIGN KEY FOR '+asm'= '7c57ab8da650a18587feaf467d5728'
CellCLI> ASSIGN KEY FOR '+asm'= 'ed63f41779c262dd34a00c0d83590b8',
```

---

**Chapter 8**

**CellCLI Command Reference**

8-68
Example 8-69  Removing Keys from Clients

This example shows how to use the ASSIGN KEY command to remove keys from clients.

CellCLI> ASSIGN KEY FOR CELL '4839deff903625aabb394df7638e7b29a'

CellCLI> ASSIGN KEY FOR ASMCLUSTER
asm1='118af47c57ab8da650ab67d5587fe728'

8.7.3 CALIBRATE

Purpose

The CALIBRATE command runs raw performance tests on cell disks, enabling you to verify the disk performance before the cell is put online.

Syntax

CALIBRATE [FLASHDISK | HARDDISK | LUN1 [, LUN2]] [FORCE]

Usage Notes

You must be logged on to the cell as the root user to run CALIBRATE.

The FORCE option enables you to run the tests when Cell Server is running. If you do not use the FORCE option, then Cell Server must be shut down. Running CALIBRATE at the same time as a Cell Server process impacts performance.

Use the FLASHDISK option to specify that only flash LUNs be calibrated.
Use the `HARDDISK` option to specify that only hard disk LUNs be calibrated.

Use the `LUNn` option to specify a list of LUNs by name be calibrated.

Examples

**Example 8-70  Output from CALIBRATE Command with FORCE Option on Oracle Exadata Storage Server**

This example shows the output when using `CALIBRATE` with `FORCE` option on Oracle Exadata Storage Server.

```
CellCLI> CALIBRATE FORCE
Calibration will take a few minutes...
Aggregate random read throughput across all hard disk luns: 1604 MBPS
Aggregate random read throughput across all flash disk luns: 4242.9 MBPS
Aggregate random read IOs per second (IOPS) across all hard disk luns: 4927
Aggregate random read IOs per second (IOPS) across all flash disk luns: 148695
Controller read throughput: 1608.05 MBPS
Calibrating hard disks (read only) ...
<table>
<thead>
<tr>
<th>Lun</th>
<th>Drive</th>
<th>Random Read Throughput</th>
<th>IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0_0</td>
<td>[20:0]</td>
<td>153.41 MBPS, and 412 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_1</td>
<td>[20:1]</td>
<td>155.38 MBPS, and 407 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_10</td>
<td>[20:10]</td>
<td>151.24 MBPS, and 423 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_11</td>
<td>[20:11]</td>
<td>152.70 MBPS, and 422 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_2</td>
<td>[20:2]</td>
<td>155.42 MBPS, and 423 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_3</td>
<td>[20:3]</td>
<td>153.14 MBPS, and 428 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_4</td>
<td>[20:4]</td>
<td>154.06 MBPS, and 424 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_5</td>
<td>[20:5]</td>
<td>150.82 MBPS, and 409 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_6</td>
<td>[20:6]</td>
<td>154.61 MBPS, and 426 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_7</td>
<td>[20:7]</td>
<td>154.46 MBPS, and 424 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_8</td>
<td>[20:8]</td>
<td>154.63 MBPS, and 426 IOPS</td>
<td></td>
</tr>
<tr>
<td>0_9</td>
<td>[20:9]</td>
<td>154.63 MBPS, and 426 IOPS</td>
<td></td>
</tr>
</tbody>
</table>
Calibrating flash disks (read only, note that writes will be significantly slower) ...
<table>
<thead>
<tr>
<th>Lun</th>
<th>Drive</th>
<th>Random Read Throughput</th>
<th>IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1_0</td>
<td>[10:0:0]</td>
<td>269.11 MBPS, and 19635 IOPS</td>
<td></td>
</tr>
<tr>
<td>1_1</td>
<td>[10:0:1]</td>
<td>268.86 MBPS, and 19640 IOPS</td>
<td></td>
</tr>
<tr>
<td>1_2</td>
<td>[10:0:2]</td>
<td>268.68 MBPS, and 19645 IOPS</td>
<td></td>
</tr>
<tr>
<td>1_3</td>
<td>[10:0:3]</td>
<td>268.92 MBPS, and 19640 IOPS</td>
<td></td>
</tr>
</tbody>
</table>
```
Lun 2_0 on drive [[12:0:0:0]] random read throughput: 269.78 MBPS, and 20436 IOPS
Lun 2_1 on drive [[12:0:1:0]] random read throughput: 269.69 MBPS, and 20394 IOPS
Lun 2_2 on drive [[12:0:2:0]] random read throughput: 269.04 MBPS, and 20439 IOPS
Lun 2_3 on drive [[12:0:3:0]] random read throughput: 269.51 MBPS, and 20420 IOPS
Lun 4_0 on drive [[9:0:0:0]] random read throughput: 269.07 MBPS, and 19668 IOPS
Lun 4_1 on drive [[9:0:1:0]] random read throughput: 269.24 MBPS, and 19697 IOPS
Lun 4_2 on drive [[9:0:2:0]] random read throughput: 269.09 MBPS, and 19676 IOPS
Lun 4_3 on drive [[9:0:3:0]] random read throughput: 269.03 MBPS, and 19681 IOPS
Lun 5_0 on drive [[11:0:0:0]] random read throughput: 268.06 MBPS, and 19714 IOPS
Lun 5_1 on drive [[11:0:1:0]] random read throughput: 268.24 MBPS, and 19696 IOPS
Lun 5_2 on drive [[11:0:2:0]] random read throughput: 268.33 MBPS, and 19717 IOPS
Lun 5_3 on drive [[11:0:3:0]] random read throughput: 268.14 MBPS, and 19693 IOPS
CALIBRATE results are within an acceptable range.

CALIBRATE stress test is now running...
Calibration has finished.

**Example 8-71  Calibrating LUNs by Name**

CALIBRATE '2_1', '2_3' FORCE

**Related Topics**

- ALTER CELL

### 8.7.4 CREATE

**Purpose**

The CREATE command creates a new object and assigns initial attributes to the object.

**Syntax**

```
CREATE object_type [name] ...
  [attribute_name = attribute_value [, attribute_name = attribute_value]...]
```

**Usage Notes**

- When multiple objects are valid as the target of a CREATE command, there is the possibility of partial success. If an error occurs, then the command is interrupted, and the remaining objects are not created.
Related Topics

- **About CellCLI Object Types**
  There are several Oracle Exadata System Software object types that can be used with CellCLI object commands.

### 8.7.4.1 CREATE CELL

**Purpose**

The `CREATE CELL` command creates a cell object and assigns initial attributes to the object.

**Syntax**

```plaintext
CREATE CELL [name]
   interconnect1=intValue1 [, interconnect2=intValue2 ...]
   [, attributeName = attributeValue ...]
```

**Usage Notes**

- The attributes that can be set are shown as **modifiable** in Example 8-92.
- This command can be used to assign the ASR value to the `snmpSubscriber` attribute. When specifying the `snmpSubscriber` attribute, the community name cannot contain spaces or the following characters: `="\<>
- The default cell name is set to the network host name of the cell with hyphens in the network name replaced with underscores. You can display the network name with the `uname -n` command. If you change the cell name, then you must choose a unique cell name.
- One to four interconnects can be specified. The `interconnect1` attribute must be specified if the `interconnect2` attribute is specified. The `interconnect1` and
Interconnect attributes must be specified if interconnect is specified, and so on.

- Interconnects for InfiniBand Network Fabric are `ibN`. Interconnects for RoCE Network Fabric are `reN`.

- By default, the CREATE CELL command creates cell disks, Exadata Smart Flash Cache, Flash Log, and PMEM cache. You can specify `FLASHCACHE=0` to bypass creation of flash cell disks and Exadata Smart Flash Cache on the cell disks. A non-zero value for Exadata Smart Flash Cache is taken as the total size for creating Exadata Smart Flash Cache. The size is divided evenly across the flash LUNs.

- The CREATE CELL command creates Exadata Smart Flash Log on the cell disks. Users can specify `FLASHLOG=0` to bypass creation of flash logs on the cell disks. A non-zero value for Exadata Smart Flash Log is taken as the total size for creating Exadata Smart Flash Log files. The size is divided evenly across the flash LUNs.

- The CREATE CELL eighthRack command enables or disables an Eighth Rack configuration on Oracle Exadata Database Machine X3-2 Quarter Racks or later. The options are `true` to enable the Eighth Rack configuration, and `false` to disable the Eighth Rack configuration. The CREATE CELL eighthRack=true command requires that there are no cell disks because enabling the Eighth Rack configures only half of the hard disks and flash capacity.

- The CREATE CELL flashCacheCompress command creates a cell with or without flash cache compression. Flash cache compression maximizes the amount of data in cache, and improves the cache hit rate. The options are `true` to create a cell with flash cache compression enabled, and `false` to create a cell without flash cache compression. The system must be restarted in order to change flash cache compression.

Note:
Flash cache compression is only supported on Oracle Exadata Database Machine X3 or Oracle Exadata Database Machine X4 racks

Starting with Oracle Exadata System Software release 19.1.0, the `httpsAccess` attribute can be used to specify a list of IP addresses or IP subnet masks that control who can access the RESTful service via HTTPS. The value you specify for `httpsAccess` overwrites any previous value. You can use the following values for `httpsAccess`:

- `ALL` — to allow access to all hosts (Default)
- `NONE` — to disable the HTTPS port completely
- `IP1, IP2,..., IPn` — to only allow access to hosts with IP addresses IP1, IP2,..., IPn where IPn is a valid IP address in IPv4, IPv4 subnet, IPv6 or IPv4-embedded IPv6 format. You can specify a maximum of 512 IP addresses for the access control list.

Additionally, instead of a single IP address, you can use the `/` character to specify a range of IP addresses using a subnet mask. For example the range `192.168.10.0/24` corresponds to hosts having IP addresses from 192.168.10.1 to 192.168.10.255. If you specify an IP address range, you need to enclose the IP address string in quotes.
Examples

Example 8-72 Creating a Cell

This example shows how to create a cell. In the example, the interconnections are set to an existing InfiniBand Network Fabric connection.

```
CellCLI> CREATE CELL cell22 interconnect1=bondib0
```

Example 8-73 Creating an Eighth Rack Configuration

This example shows how to create a cell in an Eighth Rack configuration.

```
CellCLI> CREATE CELL eighthRack=true
```

Example 8-74 Creating a Cell with Restricted HTTPs Access

This example shows how to create a cell that allows HTTPs port access only from hosts having IP addresses in the range 192.168.10.1 to 192.168.10.255.

```
CellCLI> CREATE CELL httpsAccess='192.168.10.0/24'
```

Related Topics

- Restrictions on Values of Common Attributes
  Review the following restrictions for the values of attributes used by multiple CellCLI objects.
- Oracle Exadata Database Machine Maintenance Guide

8.7.4.2 CREATE CELLDISK

Purpose

The `CREATE CELLDISK` command creates a cell disk object and assigns initial attributes to the object. You can use the `ALL` option to automatically create cell disks.

Syntax

```
CREATE CELLDISK  
   {  ALL [FLASHDISK | HARDDISK]  
      | cdisk1 attribute_name=value,[attribute_name=value]... [FORCE]  
      | ((name=cdisk2,attribute_name=value,[attribute_name=value]...)  
      |,[,(name=cdisk3,attributename=value,[attributename=value]...)]...)  
   }
```

Usage Notes

The attributes that can be specified during creation are the cell disk name (`cdiskN`), comment, lun, size, and physicalDisk.

- Either `lun` or `physicalDisk` is required when adding a specifically-named cell disk.
  - When a physical disk is provided, a single-disk LUN is created, and that LUN is used to create the cell disk. The LUN is flagged as automatically-created.
– When a LUN is provided, that device is used to create the cell disk.

• You can use the size attribute when adding a new disk that is a different size than the existing disks. You must specify a value for size that is less than or equal to the maximum allowed cell disk size.

• CREATE CELLDISK ALL is a shortcut command to create all candidate cell disks for the cell. This operation occurs in two steps:
  – All physical disks that are not already configured as LUNs are configured as single-disk LUNs (SDLs). These LUNs are flagged as automatically-created LUNs.
  – All LUNs that are not configured as cell disks are used to create cell disks. These cell disks are initially named according to the template CD_lunname_cellname. This name can later be changed using the ALTER CELLDISK command.

Note:

LUNs with a second or third extended file system (ext2/ext3) are ignored during the CREATE CELLDISK ALL operation.

– You can include size to create all cell disks with the specified size, instead of attempting to use the entire disk. If you do not specify size with the CREATE CELLDISK ALL command, then:
  * On data disks, Management Server (MS) creates the cell disks with a size that is equal to the minimum physical disks size across all physical disks in a cell (which is also the maximum allowed cell disk size in a cell).
  * On system disks, MS creates the cell disks with a size that is equal to the maximum allowed cell disk size-size of system partitions.

• The FLASHDISK option limits the CREATE CELLDISK command to cell disks that are flash disks.

• The HARDDISK option limits the CREATE CELLDISK command to cell disks that are hard disks.

• The list form of CREATE CELLDISK enables you to add a series of cell disks in a single command.

• The FORCE keyword overrides the following error conditions:
  – The physical disk provided is already part of an existing LUN.
  – The LUN provided is already associated with a cell disk.

  FORCE causes the LUN to be reused to create the new cell disk. Any preexisting configuration is lost. FORCE is not an option for CREATE CELLDISK ALL or for the list form of CREATE CELLDISK.

• The INTERLEAVING option has been deprecated. Starting with Oracle Exadata System Software release 19.1.0, attempts to create interleaving grid disks will be automatically converted to normal grid disk creation. Interleaving grid disks created in earlier Oracle Exadata System Software releases will continue to operate normally.
Cell disks are created automatically using the CREATE CELLDISK ALL command. This command creates single-disk LUNs from all available physical disks, and then creates cell disks from all available LUNs.

When a cell disk is created, metadata describing the cell disk is written to the cell disk itself and to the cell configuration files. Approximately 48 MB of the cell disk is allocated for the cell disk metadata partition. On a subsequent restart, Cell Server (CELLSRV) attempts to rediscover the created cell disk by reading configuration data on the disk. Any hardware changes in the cell might cause a change in the LUN and device name for a cell disk. The rediscovery mechanism that runs during the cell restart process changes the cell disk configuration accordingly.

**Example 8-75  Creating a Cell Disk**

```
CellCLI> CREATE CELLDISK ALL

CellCLI> CREATE CELLDISK cdisk03 lun=0_3

CellCLI> CREATE CELLDISK cdisk04 physicalDisk='I2:1:2'

CellCLI> CREATE CELLDISK CD_08_cell06 lun=0_8, size=300M

CellCLI> CREATE CELLDISK CD_03_cell04 lun=0_3

CellCLI> CREATE CELLDISK CD_05_cell09 physicalDisk='I2:1:2'
```

**Related Topics**

- ALTER CELL
- Restrictions on Values of Common Attributes  
  Review the following restrictions for the values of attributes used by multiple CellCLI objects.
- Create the Storage Cell Disks  
  Use the CREATE CELLDISK command to create the cell disks.

### 8.7.4.3 CREATE DIAGPACK

**Purpose**

The CREATE DIAGPACK command creates a diagnostic package, which contains logs and traces that you can use to troubleshoot problems in your system. You can also send the generated package to Oracle Support Services, as needed.

**Syntax**

```
CREATE DIAGPACK packStartTime=time, [durationInHrs=duration]
```

or

```
CREATE DIAGPACK alertName=alertName
```
Usage Notes

When an alert occurs, a diagnostic package is created automatically. This package contains logs and traces related to the alert.

The **CREATE DIAGPACK** command enables you to generate diagnostic packages manually.

- **The** `packStartTime` parameter specifies when to start collecting the logs and traces. The format of `packStartTime` is `yyyy_MM_ddTHH_mm_ss`, For example: `2015_07_07T09_00_00`.
  
  You can also specify the keyword `now` for `packStartTime`. The `packStartTime` cannot be in the future and cannot be older than 7 days. The value of `packStartTime` is used as part of the name of the diagnostic package.

- **The** `durationInHrs` parameter specifies the number of hours of logs and traces to include in the diagnostic package. Valid values are from 1 (default) to 6.
  
  Every diagnostic package includes logs 1 hour before and 1 hour after the `packStartTime`. For example, if you specify a time of `12_00_00`, then logs will collected from `11_00_00` to `13_00_00`, unless the end time is in the future.

- **The** `alertName` parameter specifies the alert name for which to create the diagnostic package. You can run the `LIST ALERTHISTORY` command to view the alert names.

Name of Diagnostic Packages

The name of the diagnostic package is formed as:

`hostname_diag_packStartTime_unique package ID`

For example: `testcell1_diag_2015_07_07T09_00_00_3`

For alerts, the name of the diagnostic package is formed as:

`hostname_timestamp of when the package was created_alert ID`. For example:

`testcell1_2015_09_30T13_13_00_2_1`

Location of Diagnostic Packages

The location of the diagnostic packages is `$LOG_HOME`.

Status of Diagnostic Packages

You can run the `LIST DIAGPACK` command to get a list of diagnostic packages in your system, and their status.

Privileges Needed to Create, List, and Download Diagnostic Packages

There are certain privileges that are needed for working with diagnostic packages. Use CellCLI to grant the following privileges to a role:

- Privilege to create diagnostic packages:
  
  ```
  grant privilege CREATE ON DIAGPACK to ROLE role
  ```
Privilege to list diagnostic packages and check their status:

grant privilege LIST ON DIAGPACK to ROLE role

Privilege to download diagnostic packages:

grant privilege DOWNLOAD ON DIAGPACK to ROLE role

You can then grant the role to users. For example, if you named your role diagpack_role, the following command grants the role to fred.

CellCLI> GRANT ROLE diagpack_role TO USER fred

During deployment, Oracle Exadata Deployment Assistant (OEDA) creates an Exadata storage software user called CELLDIAG. You can use this user to connect to a cell remotely using ExaCLI or REST API. This user has privileges to create, list, and download diagnostic packages.

**Downloading Diagnostic Packages**

You can download diagnostic packages using any of the following methods. Note that you need the DOWNLOAD ON DIAGPACK privilege before you can download diagnostic packages.

- **Using the REST API**
  
  To download the diagnostic package by name, use the following URL, where
  * hostname specifies the host name of the cell and diagpackname specifies the name of the diagnostic package:

  https://hostname/diagpack/download?name=diagpackname

  If the user is not already logged in, the URL will prompt for a user name and password.

  Diagnostic packages can also be accessed at https://hostname/diagpack. For example: https://cell1.example.com/diagpack.

  The page then prompts the user to log in:

  User: fred
  Password: ********

  Based on the user’s privileges, various sections of this page could be hidden:

  * The form to create a new diagpack will not be shown if the user does not have the CREATE ON DIAGPACK privilege.

  * Similarly, the list of alerts and their diagnostic packages will not be shown if the user does not have the LIST ON DIAGPACK privilege.
To download the diagnostic package by alert name, use the following URL, where hostname specifies the host name of the cell and alertName specifies the alert name of the diagnostic package:

https://hostname/diagpack/download?alert=alertName

The alert name is the same alert name that is used in AlertHistory. It looks like 1, 2, 3 for stateless alerts, and 1_1, 2_1, 3_1, 3_2 for stateful alerts.

**Using the download ExaCLI command**

ExaCLI enables you to run CellCLI commands on storage nodes remotely from compute nodes.

To run the download command, run the following commands on a compute node:

1. **Start up ExaCLI and connect to the cell containing the diagnostic pack.** For example, use a command similar to the following where hostname specifies the host name of the cell:

   ```
   exacli -l celladministrator -c hostname
   Password=********
   ```

2. **Run the download command** using a command similar to the following where name specifies the name of the diagnostic package to download and destinationFolder specifies the directory where you want to save the downloaded diagnostic package:

   ```
   exacli> download diagpack name destinationFolder
   ```

**Getting the diagnostic package from the alert emails**

The alert emails include diagnostic packages for all alerts except INFO, CLEAR, and WARNING. Diagnostic packages are generated for critical alerts only.

**Re-triggering Package Creation from the Web Page**

You can use the following URL to re-trigger package creation:

https://hostname/diagpack

If the diagnostic package for an alert does not exist on disk, then the web page shows a Create Package link instead of a Download link.

Click the Create Package link to add the alert to the list for creating a diagnostic package. Once the diagnostic package has been created, and the web page is refreshed, the page will display a Download link that you can use to download the newly created diagnostic package.

**Turning Off the Diagnostic Pack Attachment in Emails**

To turn off the diagnostic pack attachment in emails, run **ALTER CELL diagPackEmailAttach=FALSE**. The diagnostic packs are still generated and stored on the storage servers. To download the diagnostic packs, see "Downloading Diagnostic Packages".
Examples

Example 8-76  Using "now" for packStartTime

This example creates a diagnostic package using now as the start time and the default duration of 1 hour.

The output is 1 compressed file under $LOG_HOME.

CellCLI> CREATE DIAGPACK packStartTime="now"
Processing: scab01cell11_diag_2015_07_08T17_53_58_1
Use 'list diagPack' to check its status.

Example 8-77  Specifying a duration

This example creates 3 diagnostic packages under $LOG_HOME:

The first package has a start time of 2015_07_07T09_00_00.

The second package has a start time of 2015_07_07T10_00_00.

The third package has a start time of 2015_07_07T11_00_00.

CellCLI> CREATE DIAGPACK packStartTime="2015_07_07T09_00_00", durationInHrs=3
   Processing: scab01cell11_diag_2015_07_07T09_00_00_1
   scab01cell11_diag_2015_07_07T10_00_00_1 (In queue...)
   scab01cell11_diag_2015_07_07T11_00_00_1 (In queue...)
   Use 'list diagPack' to check its status.

Related Topics

• About Oracle Auto Service Request
• About Automatic Diagnostic Repository
• Using the ExaCLI Utility

8.7.4.4 CREATE FLASHCACHE

Purpose

The CREATE FLASHCACHE command creates Exadata Smart Flash Cache on a cell for I/O requests.

Syntax

CREATE FLASHCACHE { ALL [size=fc_size] | CELLDISK="cdisk1 [,cdisk2] ..." [, size=fc_size] }

Usage Notes

Cell disks defined on Exadata Smart Flash Cache cannot be exported.

The ALL argument creates Exadata Smart Flash Cache on all flash cell disks. If the ALL argument is not specified, then the CELLDISK argument must be specified.
Using the `CELLDISK` argument, you can specify a list of flash cell disks to be used for flash cache. The names of the flash cell disks are comma-delimited. The `FLASHDISK` argument is not required.

The `size` argument specifies the total space used for flash cache. Similar to space in grid disks and flash logs, flash cache space is allocated in 16 MB units, referred to as allocation units. If the `size` attribute is specified when creating a flash cache, then the size allocated is the size of the largest multiple of allocation units less than or equal to the specified size. For example, if `300M` is specified for the `size` attribute, then 288 MB (16x18) is allocated because 288 is the largest multiple of 16 that is less than or equal to 300.

A minimum of 1 allocation unit is always allocated, so the minimum size for a flash cache is 16 MB. Any size value less than 16 MB is rounded up to 16 MB.

Before specifying the `size` attribute, ensure that you have first determined the available free space on each target flash cell disk with the `LIST FLASHCACHE` command. For example, `LIST FLASHCACHE ATTRIBUTES freespace`. If the `size` attribute is not specified, then the maximum size is allocated.

If the `size` attribute is not specified, then all available space on each cell disk in the list is used for Exadata Smart Flash Cache.

By default, 5 percent of space on Extreme Flash storage servers is used for write-back flash cache.

**Examples**

**Example 8-78 Creating Exadata Smart Flash Cache**

This example shows how to create Exadata Smart Flash Cache on a cell.

```
CellCLI> CREATE FLASHCACHE ALL
CellCLI> CREATE FLASHCACHE ALL SIZE=250g
CellCLI> CREATE FLASHCACHE CELLDISK='fd_01,fd_02,fd_03,fd_04'
CellCLI> CREATE FLASHCACHE CELLDISK='fd_01_mycell,fd_02_mycell', size = 64G
```

**Related Topics**

- `ALTER FLASHCACHE`

**8.7.4.5 CREATE FLASHLOG**

**Purpose**

The `CREATE FLASHLOG` command creates the Oracle Exadata Smart Flash Log on a cell for redo log I/O requests.
Syntax

CREATE FLASHLOG {{ALL {{FLASHDISK} | [size=log_size]} | CELLDISK='cdisk1 \
[,cdisk2]...'} [, size=log_size]}

Usage Notes

The `CREATE FLASHLOG` command accepts a list of comma-delimited flash cell disks. If a size is specified in the command, then that size is divided evenly across the cell disks, and will total the specified size. If a size is not specified, then a default size of 512 MB is used.

The size of Oracle Exadata Smart Flash Log space on each flash disk must be less than 4 GB. If all 16 flash disks are available, then the total size of Oracle Exadata Smart Flash Log must be less than 64 GB.

Similar to space in grid disks and flash cache, flash log space is allocated in 16 MB units, referred to as allocation units. If the size attribute is specified when creating a flash log, then the size allocated is the size of the largest multiple of allocation units less than or equal to the specified size. For example, if 300M is specified for the size attribute, then 288 MB (16x18) is allocated because 288 is the largest multiple of 16 that is less than or equal to 300.

A minimum of 1 allocation unit is always allocated, so the minimum size for a flash log is 16 MB. Any size value less than 16 MB is rounded up to 16 MB.

The `ALL FLASHDISK` argument creates Oracle Exadata Smart Flash Log on all flash cell disks. If the `ALL` argument is not specified, then the `CELLDISK` argument must be specified. The `FLASHDISK` argument is not required.

Note:
The `CREATE FLASHCACHE` command, by default, uses all available space on each flash disk. Therefore, use the `CREATE FLASHLOG` command before creating the flash cache to ensure both objects consume the correct amount of flash disk space.

To change the size of the flash log, use the `DROP FLASHLOG` command to drop the flash log, and then use the `CREATE FLASHLOG` command to create a flash log with the new size.

Examples

The following example shows how to create Oracle Exadata Smart Flash Log on a cell.

**Example 8-79  Creating Oracle Exadata Smart Flash Log**

CellCLI> CREATE FLASHLOG ALL

CellCLI> CREATE FLASHLOG ALL SIZE=1g
CellCLI> CREATE FLASHLOG ALL FLASHDISK
CellCLI> CREATE FLASHLOG CELLDISK='fd1,fd2,fd3,fd4'

8.7.4.6 CREATE GRIDDISK

Purpose

The CREATE GRIDDISK command creates a grid disk object on a specified cell disk or creates one grid disk on each cell disk on the cell. The command also assigns initial attributes to the new grid disks.

Syntax

CREATE GRIDDISK { ALL [FLASHDISK | HARDDISK] PREFIX=gdisk_name_prefix
| gdisk CELLDISK = attribute_value 
[, attribute_name = attribute_value]... }

Usage Notes

The attributes that can be specified are the grid disk name (gdisk), CELLDISK, size, and comment. The CELLDISK argument is required when a single grid disk is created.

• If an individual name is entered with the command, then the grid disk is created on the cell disk specified by the CELLDISK argument. You must ensure that the grid disk name is unique across all cells. If the disk name is not unique, then it might not be possible to add the grid disk to an Oracle ASM disk group.

• The length of a grid disk name is limited to 30 characters.

• The FLASHDISK option limits the CREATE GRIDDISK command to cell disks that are flash disks.

• The HARDDISK option limits the CREATE GRIDDISK command to cell disks that are hard disks.

• If the ALL PREFIX option is entered with the command, then one grid disk is created on each cell disk on the cell. PREFIX is required when ALL is used.

The PREFIX option specifies the prefix assigned to the names of the created grid disks. The generated grid disk names are composed of the grid disk prefix followed by an underscore (_) and then the cell disk name in the following form:

gdisk_name_prefix_cdisk_name

Choose a prefix for the grid disk name that matches the Oracle ASM disk group to which the grid disk belongs to help you identify which disks belong to a disk group. The generated grid disk (gdisk_name_prefix_cdisk_name) must follow the restrictions on the name value.

For example, if the Oracle ASM disk group name is data01, then data01 is used as the prefix for the grid disk names. If CREATE GRIDDISK ALL PREFIX=data01 is run on a cell with cell disks CD_01_cell101, CD_02_cell101, and CD_03_cell101, then grid disks are named data01_CD_01_cell101, data01_CD_02_cell101, and data01_CD_03_cell101 are created on each cell disk respectively.
The CREATE GRIDDISK ALL command skips disks which do not have enough free space for a minimum grid disk. A message stating which disks did not have enough free space appears, and the command continues.

The size and offset attributes are optional attributes specified as a number in bytes, unless the suffix M (megabytes), G (gigabytes), or T (terabytes) is included with the number, such as size=300M, or size=150G.

Grid disk space is allocated in 16 MB units, referred to as allocation units. If the size attribute is specified when creating a grid disk, then the size allocated is the size of the largest multiple of allocation units less than or equal to the specified size. For example, if 300M is specified for the size attribute, then 288 MB (16x18) is allocated because 288 is the largest multiple of 16 that is less than or equal to 300.

A minimum of 1 allocation unit is always allocated, so the minimum size for a grid disk is 16 MB. Any size value less than 16 MB is rounded up to 16 MB.

Before specifying the size attribute, ensure that you have first determined the available free space on each target cell disk with the LIST CELLDISK command. For example, LIST CELLDISK cdisk ATTRIBUTES freespace. If the size attribute is not specified, then the maximum size is allocated.

Offset determines the position on the disk where the grid disk is allocated. The outermost tracks have lower offset values, and these tracks have greater speed and higher bandwidth. Offset can be explicitly specified to create grid disks that are relatively higher performing than other grid disks. If offset is not specified, then the best (warmer) available offset is chosen automatically in chronological order of grid disk creation. You should first create the grid disks that are expected to contain the most frequently accessed (hottest) data, and then create the grid disks that contain the relatively colder data.

Note:
Offset does not apply to flash storage on Extreme Flash storage servers.

The value of the availableTo attribute is set to the names of the clients that you want to set up for DB-scoped security. These clients were assigned security keys that match keys in the configuration files on the hosts.

The value entered for a client name is the unique name (DB_UNIQUE_NAME). The specified clients are those that are allowed to access the grid disk. If a value is entered for availableTo, then only the specified clients have access to the grid disk; otherwise, any client can have access.

Do not edit the value of idp.type or idp.boundary. Oracle Exadata System Software passes a hint to Oracle ASM about the type of interleaved grid disk, either normal redundancy or high redundancy. Oracle ASM sets the default value for idp.type to static and idp.boundary to the type of redundancy used in the underlying grid disks. The default value of the idp.type attribute is static for Oracle Exadata Storage Server disk groups created on interleaved grid disks.
Interleaved grid disks are deprecated in Oracle Exadata System Software release 19.1.0.

- The `cachingPolicy` attribute can be set to `default` or `none`. The `default` option allows the data to be cached in flash cache for the grid disk. The `none` option means that flash cache is not used for data for the grid disk.

  Flash cache is not used with data files placed in a disk group composed of grid disks with their `cachingPolicy` set to `none`.

  Oracle Exadata Deployment Assistant (OEDA) configures the RECO disk group to have the `cachingPolicy` of its grid disks set to `none`; therefore flash cache is not used for any data files placed in the RECO disk group.

- The `virtualSize` attribute is used to create the SPARSE disk group. The maximum virtual size for a sparse grid disk is approximately 100 TB. Sparse grid disks are available for Oracle Exadata Database Machine X3-2 and later.

The Oracle Database and Oracle Grid Infrastructure software must be release 12.1.0.2.0 BP5 or later when using sparse grid disks.

### Example 8-80  Creating a Grid Disk

This examples shows how to create grid disks.

```
CellCLI> CREATE GRIDDISK data1_CD_01_cell01 CELLDISK=CD_01_cell01, size=200G
CellCLI> CREATE GRIDDISK data2_CD_02_cell01 CELLDISK=CD_02_cell01, size=200G
CellCLI> CREATE GRIDDISK data3_CD_03_cell01 CELLDISK=CD_03_cell01
CellCLI> CREATE GRIDDISK ALL PREFIX=data01, availableTo='+asm,db1,db2'
CellCLI> CREATE GRIDDISK hr7_CD_07_cell01 CELLDISK=CD_07_cell01, availableTo='asm_hr,hrdb0'
CellCLI> CREATE GRIDDISK GD123 CELLDISK=RECO_CD123, size=100G, cachingPolicy=none
```

### Example 8-81  Creating a SPARSE Disk Group

```
CellCLI> CREATE GRIDDISK spar01 celldisk=CD_01_cell01, size=10G, virtualsize=100G
```
Related Topics

- **Restrictions on Values of Common Attributes**
  Review the following restrictions for the values of attributes used by multiple CellCLI objects.
- **ALTER GRIDDISK**

### 8.7.4.7 CREATE KEY

**Purpose**
The `CREATE KEY` command creates and displays a random hexadecimal string to assign client keys. The use of `CREATE KEY` ensures that the security key is in the correct format. This command provides a way to generate a key in the correct format, and it can be run on any cell.

**Syntax**
```plaintext
CREATE KEY
```

**Usage Notes**
The security key must be entered in the `cellkey.ora` configuration file on the computer hosts that contain clients for which you want to authorize access to a cell.

The key is also assigned to clients that access grid disk storage.

The key must be copied manually to the hosts and cells.

**Example 8-82 Creating a Key**
This example shows the `CREATE` command with the `KEY` object.

```
CellCLI> CREATE KEY
3452c64fec9a5800bbe48d4093269400
```

**Related Topics**
- About Security Keys
- **ASSIGN KEY**

### 8.7.4.8 CREATE PMEMCACHE

**Purpose**
The `CREATE PMEMCACHE` command creates a PMEM Cache for I/O requests.

**Syntax**
```plaintext
CREATE PMEMCACHE {[ALL [size = cache_size] | CELLDISK = 'cdisk1 [,cdisk2],,,']
                                              [, size=cache_size]}
```

Usage Notes

Cell disks used by PMEM Cache cannot be exported.

The ALL argument creates the PMEM Cache on all PMEM cell disks. If the ALL argument is not specified, then the CELLDISK argument must be specified.

Using the CELLDISK argument, you can specify a list of PMEM cell disks to be used for PMEM cache. cdiskn represents a cell disk name. The names of the PMEM cell disks are comma-delimited and enclosed in quotation marks.

The size argument specifies the total size of the PMEM cache. Similar to space in grid disks, PMEM cache space is allocated in 16 MB units, referred to as allocation units. If the size attribute is specified when creating a PMEM cache, then the size allocated is the size of the largest multiple of allocation units less than or equal to the specified size. For example, if 300M is specified for the size attribute, then 288 MB (16x18) is allocated because 288 is the largest multiple of 16 that is less than or equal to 300.

A minimum of 1 allocation unit is always allocated, so the minimum size for a PMEM cache is 16 MB. If you specify a size less than 16 MB, then it is rounded up to 16 MB.

Before specifying the size attribute, ensure that you have first determined the available free space on each target PMEM cell disk with the LIST CELLDISK ATTRIBUTES freespace WHERE disktype=PMEM command.

If the size attribute is not specified, then the maximum size is allocated. All available space on each cell disk in the list is used for PMEM Cache.

Examples

Example 8-83 Creating PMEM Cache

The following command demonstrates how to create a 64 GB PMEM cache on specific PMEM cell disks.

CREATE PMEMCACHE celldisk='PM_01_mycell,PM_02_mycell', size = 64G

The following command demonstrates how to create a 64 GB PMEM cache using all PMEM cell disks.

CREATE PMEMCACHE ALL size = 64G

8.7.4.9 CREATE PMEMLOG

Purpose

The CREATE PMEMLOG command creates the PMEM log file on a cell for redo log I/O requests.

Syntax

CREATE PMEMLOG {ALL [size=log_size] | CELLDISK='cdisk1[,cdisk2]...' [, size=log_size]}
Usage Notes

The `CREATE PMEMLOG` command accepts a list of comma-delimited PMEM cell disks. If a size is specified in the command, then that size is divided evenly across the PMEM cell disks. Each PMEM cell disk must be at least 16 MB in size. If a size is not specified, then a default size of 960 MB is used.

The `ALL` argument creates `PMEMLOG` on all PMEM cell disks. If the `ALL` argument is not specified, then the `CELLDISK` argument must be specified.

![Note:]

The `CREATE PMEMCACHE` command, by default, uses all available space on each PMEM cell disk. Therefore, use the `CREATE PMEMLOG` command before creating the PMEM cache to ensure both objects consume the correct amount of PMEM disk space.

To change the size of the PMEM log, use the `DROP PMEMLOG` command to drop the PMEM log, and then use the `CREATE PMEMLOG` command to create a PMEM log with the new size.

Examples

The following example shows how to create PMEM log file on a cell.

**Example 8-84  Creating the PMEM log file**

To create the PMEMLOG using the default size that uses all available PMEM cell disks, use the following command:

`CellCLI> CREATE PMEMLOG ALL`

To create the PMEMLOG of size 1 GB that uses all available PMEM cell disks, use the following command:

`CellCLI> CREATE PMEMLOG ALL SIZE=1g`

To create the PMEMLOG on specific PMEM cell disks, use the following command:

`CellCLI> CREATE PMEMLOG CELLDISK='nv1,nv2,nv3,nv4'`

To create the PMEMLOG that is 1 GB in size on specific PMEM cell disks, use the following command:

`CellCLI> CREATE PMEMLOG CELLDISK='nv1,nv2,nv3,nv4', size=1G`

8.7.4.10 CREATE QUARANTINE

Purpose

The `CREATE QUARANTINE` command allows a quarantine to be created manually.
Syntax

CREATE QUARANTINE quarantineType=value quarantinePlan="SYSTEM", 
   dbUniqueName=value[, attributename=value]...

Usage Notes

Manual creation of quarantines should be done in coordination with Oracle Support Services. In general, manual quarantines are created to proactively isolate SQL statements that are known to cause problems.

- `quarantineType` specifies the type of quarantine to be created, such as SQLID and SQL_PLAN.
- `quarantinePlan` must be set to SYSTEM. Oracle Support Services may specify other values.
- `dbUniqueName` specifies the name of the database that has the quarantine.

Example 8-85   Creating a Quarantine

This example shows the `CREATE` command with the `QUARANTINE` object.

CELLCLI> CREATE QUARANTINE  DBUG comment='For debugging quarantines'

CELLCLI> CREATE QUARANTINE quarantineType="SQLID", 
   quarantinePlan="SYSTEM", -
   dbUniqueName="DB1", sqlid="5xnjp4cutcl37";

Related Topics

- DESCRIBE QUARANTINE

8.7.4.11 CREATE ROLE

Purpose

The `CREATE ROLE` command creates a role for a user accessing a cell.

Syntax

CREATE ROLE  role_name [, role_name2, ...]

Usage Notes

The role name should be unique.

Example 8-86   Creating a Role

This example shows how to create a role named `gd_monitor`.

CellCLI>CREATE ROLE gd_monitor
8.7.4.12 CREATE THRESHOLD

Purpose

The CREATE THRESHOLD command creates a threshold object that specifies the conditions for generation of a metric alert.

Syntax

CREATE THRESHOLD name attributename=value [, attributename=value]...

Usage Notes

The attributes that can be specified are comparison, critical, occurrences, observation, and warning.

- The name argument is required. The name is comprised of a metric name and an object name with the format metricName.objectName, such as db_io_rq_sm_sec.db123 or ct_io_wt_rq.interactive. Use the LIST METRICCURRENT metric command to display the available object name for metric. The object name is optional.
- When a object name is not specified, then the threshold is applied to all metric objects for the given metric.
- The comparison attribute is required with a condition value. The value must be '<', '<=', '=', '>=', or '>'.
- The occurrences attribute specifies the number of consecutive measurements over the threshold value that trigger a state change.
- The observation attribute is the number of measurements over which measured values are averaged.
- A state change to the value set in warning or critical causes a stateful alert to be generated.
- The GD_SP_PRCT_ALLOCATED metric has a built-in threshold, and automatically sends alerts. Create thresholds for other metrics to receive alerts for those metrics.

When specifying occurrences and observations, you need the specified number of consecutive occurrences of sample averages over the number of observations to cause an alert. For example, if the following five observations (observations=5) happen on a cell, then the average sample would be 10 because the number of consecutive occurrences (occurrences=2) had values of 5 and 15.

Observation 1: 0
Observation 2: 30
Observation 3: 0
Observation 4: 5
Observation 5: 15
Example 8-87  Creating a Threshold

This example shows how to create a threshold.

CellCLI> CREATE THRESHOLD db_io_rq_sm_sec.db123 comparison='>',
critical=120

CellCLI> CREATE THRESHOLD ct_io_wt_sm.interactive warning=10,
critical=20, -
    comparison='=', occurrences=2, observation=5

Related Topics

- Exadata Alerts
  Alerts draw attention to potential and actual problems and other events of interest
to an administrator.

8.7.4.13 CREATE USER

Purpose

The CREATE USER command creates a user.

Syntax

CREATE USER name PASSWORD = *

Usage Notes

- The user name should be unique.
- `celladmin`, `cellmonitor`, and `root` are reserved user names that cannot be used
  with the CREATE USER command.
- The system prompts for a password for the new user. The password must have 12
to 40 alphanumeric characters or special characters (!@#$%^&*() with at least one
digit, one lowercase letter, and one uppercase letter. Starting with Oracle Exadata
System Software release 18.1.0.0.0, the password can be 8 to 40 characters in
length and can also utilize the special characters - and _.
- The new password cannot be the same as the current password for the user.

Example 8-88  Creating a User

This example shows how to create a user.

CellCLI> CREATE USER agarcia PASSWORD = *
password:
Confirm password: password
User agarcia successfully created.
8.7.5 DESCRIBE

Purpose

The DESCRIBE command displays a list of attributes for the object type that is provided as an argument. The list of attributes indicates whether each attribute can be modified.

Syntax

DESCRIBE object_type

Usage Notes

• The object_type is one of the object types supported by CellCLI.
• The list of attributes can be used as arguments in the LIST command.
• DESCRIBE does not display all of the attributes for the objects.

• DESCRIBE ACTIVEREQUEST
• DESCRIBE ALERTDEFINITION
• DESCRIBE ALERTHISTORY
• DESCRIBE CELL
• DESCRIBE CELLDISK
• DESCRIBE DATABASE
• DESCRIBE DISKMAP
• DESCRIBE FLASHCACHE
• DESCRIBE FLASHCACHECONTENT
• DESCRIBE FLASHLOG
• DESCRIBE GRIDDISK
• DESCRIBE IBPORT
• DESCRIBE IORMPLAN
• DESCRIBE KEY
• DESCRIBE LUN
• DESCRIBE METRICCURRENT
• DESCRIBE METRICDEFINITION
• DESCRIBE METRICHISTORY
• DESCRIBE OFFLOADGROUP
• DESCRIBE PHYSICALDISK
• DESCRIBE PLUGGABLEDATABASE
• DESCRIBE PMEMCACHE
• DESCRIBE PMEMLOG
• DESCRIBE QUARANTINE
8.7.5.1 DESCRIBE ACTIVEREQUEST

Purpose

The DESCRIBE ACTIVEREQUEST command displays a list of attributes for the ACTIVEREQUEST object type.

Syntax

DESCRIBE ACTIVEREQUEST

Usage Notes

The attributes for the DESCRIBE ACTIVEREQUEST command can include the following:

- `asmDiskGroupNumber`: Number of the Oracle ASM disk group
- `asmFileIncarnation`: Incarnation number of the Oracle ASM file
- `asmFileNumber`: Number of the Oracle ASM file
- `consumerGroupId`: Identifier of the consumer group
- `consumerGroupName`: Name of the consumer group
- `dbId`: Database unique name
- `dbName`: Database name
- `dbRequestId`: Identifier of the database request
- `fileType`: File type associated with the request
- `id`: Unique identifier of the active request
- `instanceNumber`: Instance number associated with the request
- `ioBytes`: Number of bytes of I/O against the grid disk in the current session
- `ioBytesSoFar`: Number of total bytes of I/O
- `ioGridDisk`: Grid disk used by a request
- `ioOffset`: Measure of the offset on the grid disk
- `ioReason`: Reason for I/O activity, such as a control-file read
- `ioType`: Type of active request, such as file initialization, read, write, predicate pushing, filtered backup read, or predicate push read
- `name`: Unique name of the active request
- `objectNumber`: Object number associated with the request
- `parentId`: Identifier of the parent request
• **pdbID**: Identifier of the pluggable database

• **requestState**: State of the active request, such as
  – Accessing Disk
  – Computing Result
  – Network Receive
  – Network Send
  – Queued Extent
  – Queued for Disk
  – Queued for File Initialization
  – Queued for Filtered Backup Read
  – Queued for Network Send
  – Queued for Predicate Pushing
  – Queued for Read
  – Queued for Write
  – Queued in Resource Manager

• **sessionID**: Identifier of the session

• **sessionSerNumber**: Serial number of the database session

• **sqlID**: Identifier of the SQL command associated with the request

• **tableSpaceNumber**: Tablespace number associated with the request

**Example 8-89  Describing the ACTIVEREQUEST Object**

This example shows the **DESCRIBE** command with the **ACTIVEREQUEST** object.

CellCLI> **DESCRIBE ACTIVEREQUEST**

```
name
asmDiskGroupNumber
asmFileIncarnation
asmFileNumber
consumerGroupID
consumerGroupName
dbID
dbName
dbRequestID
fileType
id
instanceNumber
ioBytes
ioBytesSofar
ioGridDisk
ioOffset
ioReason
ioType
objectNumber
parentID
```
8.7.5.2 DESCRIBE ALERTDEFINITION

Purpose

The DESCRIBE ALERTDEFINITION command displays a list of attributes for the ALERTDEFINITION object type.

Syntax

DESCRIBE ALERTDEFINITION

Usage Notes

The attributes for the DESCRIBE ALERTDEFINITION command can include the following:

• alertShortName: Abbreviated name for the alert. If the alert is based on a metric, then the short name is the same as the corresponding metric name attribute.

• alertSource: Source of the alert, such as BMC or ADR

• alertType: Type of the alert. Values are stateful or stateless.
  – Stateful alerts are automatically cleared on transition to normal.
  – Stateless alerts are never cleared. You can change the alert by setting the examinedBy attribute.

• description: Description for the alert

• metricName: Metric name if the alert is based on a metric

• name: Identifier for the alert

Example 8-90   Describing the ALERTDEFINITION Object

CellCLI> DESCRIBE ALERTDEFINITION

name
alertShortName
alertSource
alertType
description
metricName

Related Topics

• Displaying Alert Definitions
  Use the LIST ALERTDEFINITION command to display the alert definitions for the storage server.
### 8.7.5.3 DESCRIBE ALERTHISTORY

#### Purpose

The `DESCRIBE ALERTHISTORY` command displays a list of attributes for the `ALERTHISTORY` object type.

#### Syntax

```
DESCRIBE ALERTHISTORY
```

#### Usage Notes

The attributes for the `DESCRIBE ALERTHISTORY` command can include the following:

- **alertAction**: Recommended action to perform for this alert
- **alertDescription**: Description for the alert
- **alertMessage**: Brief explanation of the alert
- **alertSequenceID**: Unique sequence ID for the alert. When an alert changes its state, such as warning to critical or critical to clear, another occurrence of the alert is created with the same sequence number and a time stamp of the transition.
- **alertShortName**: Abbreviated name for the alert. If the alert is based on a metric, then the short name is the same as the corresponding metric `name` attribute.
- **alertType**: Type of the alert. Values are stateful or stateless.
  - Stateful alerts are automatically cleared on transition to normal.
  - Stateless alerts are never cleared. You can change the alert by setting the `examinedBy` attribute.
- **beginTime**: Time stamp when an alert changes its state
- **endTime**: Time stamp for the end of the period when an alert changes its state
- **examinedBy**: Administrator who reviewed the alert
- **failedMail**: Intended e-mail recipient when a notification failed
- **failedSNMP**: Intended SNMP subscriber when a notification failed
- **metricObjectName**: Object, such as cell disk or grid disk, for which a metric threshold has caused an alert
- **metricValue**: Value of the metric that caused the alert
- **name**: Unique identifier for the alert
- **notificationState**: Number indicating progress in notifying subscribers to alert messages:
  - 0, or "non-deliverable": Never tried
  - 1, or "sent": Sent successfully
  - 2, or "attempting delivery": Retrying, up to 5 times
  - 3, or "delivery failed": Retry failed 5 times
– 4, or "creating diagpack": The diagnostic packaging and the outgoing alert email are still pending for this alert. It is used by MS to keep track of the alerts that have not been processed or are being processed. This enables MS, in the event of a restart, to continue processing alerts with state 4 and send the alert email.

• `sequenceBeginTime`: Time stamp when an alert sequence ID is first created
• `serviceRequestLink`: The URL to the service request associated with the alert
• `serviceRequestNumber`: The service request number associated with the alert
• `severity`: Severity level. Values are:
  – clear
  – info
  – warning
  – critical

Example 8-91  Describing the ALERTHISTORY Object

```
CellCLI> DESCRIBE ALERTHISTORY

name     alertAction
alertDescription  alertSequenceID
alertShortName  alertType
beginTime  endTime
examinedBy              modifiable
failedMail
failedSNMP
metricObjectName
metricValue
notificationState
sequenceBeginTime
serviceRequestLink
serviceRequestNumber
severity
```

Related Topics

• DESCRIBE METRICDEFINITION

8.7.5.4 DESCRIBE CELL

Purpose

The DESCRIBE CELL command displays a list of attributes for the CELL object type.
Syntax

DESCRIBE CELL

Usage Notes

The following list contains the attributes for the DESCRIBE CELL command.

- **accessLevelPerm**: Specifies the access level at which the cell runs by default. Value is either `remoteLoginEnabled` or `remoteLoginDisabled`.
- **accessLevelTemp**: Duration of time in which the access level is temporarily changed from the setting of `accessLevelPerm`.
- **accountLockInDays**: Number of days after a password expires before a user account is locked. Available with Oracle Exadata System Software release 19.1.0 or later.
- **bbuStatus**: Status of hard disk controller battery-backed unit (BBU).
- **cellsrvStatus**: Status of Cell Server.
- **cellVersion**: Release number of the cell software.
- **comment**: User-supplied text string.
- **cpuCount**: Number of CPUs on the cell.
- **dbPerfDataSuppress**: Specifies which databases should not have their statistics reported in Automatic Workload Repository (AWR) reports.
- **diagHistoryDays**: Number of days ADR files are retained. The default is 7 days.
- **diagPackEmailAttach**: Whether a diagpack is included as an attachment in the alert email or not. The default is true.
- **diagPackUploadEnabled**: Whether the auto diagpack upload feature is enabled or not. The default is true.
- **doNotServiceLEDStatus**: Status of the cell DoNotService LED. The value can be on or off.
- **eighthRack**: Specifies whether Oracle Exadata Database Machine Eighth Rack configuration for storage cells is enabled or disabled.
- **emailFormat**: File format for email messages. The value can be `html` or `text`.
- **emailSubscriber**: List of names that subscribe to the alert notifications.
- **events**: String for `events++` that is passed to Cell Server for debugging and trace information purposes.
- **exacliEnabled**: Specifies whether `exacli` is enabled or disabled. The default value is true (enabled).
- **fanCount**: Count of working fans and total fans, displayed as working/total.
- **fanStatus**: Status of the fan. The value can be `normal`, `warning`, or `critical`.
- **flashCacheCompress**: Setting for flash cache compression. The value can be `true` or `false`.
- **flashCacheMode**: Setting for flash cache. The value can be `writethrough` or `writeback`. The default is `writethrough`.
- **httpsAccess**: Control list of IP addresses for HTTPs port access to the Exadata RESTful Service. Available with Oracle Exadata System Software release 19.1.0 or later.

- **id**: Global unique identifier (GUID) supplied by the hardware vendor

- **interconnect[1-8]**: Interconnect1 to interconnect8 for the cell. For example: bondeth0 or bondib0.

- **interconnectCount**: Number of active InfiniBand network interconnects

- **iormBoost**: Ratio of the cumulative number of positions in the I/O queue that were skipped because of IORM scheduling to the number of I/Os that where scheduled. This ratio is calculated by sampling the changes each minute in the two numbers.

- **IOTimeoutThreshold**: Specifies the timeout threshold. If cell I/O takes longer than the defined threshold, then the I/O is canceled, and Oracle ASM redirects the I/O to another mirror copy of the data.

- **ipaddress[1-8]**: ipaddress1 to ipaddress8 for the cell

- **kernelVersion**: Version of the host kernel software

- **location**: Physical location of the cell hardware supplied by the user

- **locatorLEDStatus**: Status of cell LOCATE LED. The value can be on or off.

- **makeModel**: Make and model of the cell hardware supplied by the vendor

- **memoryGB**: The memory in gigabytes for the cell

- **metricCollection**: Indicator for whether Management Server performs metric collection. Values are TRUE or FALSE. If set to FALSE, then all collection and alert mining is stopped. The default setting is TRUE.

- **metricHistoryDays**: Number of days metric history files are retained. The default is 7 days.

- **msStatus**: Status of Management Server

- **name**: Unique name for the cell

- **notificationMethod**: Notification method for alerts. The value should be mail, snmp, none, or both mail and snmp.

- **notificationPolicy**: Indicator for severity alerts to be sent to subscribers. The value for notificationPolicy should be none or any combination of critical, warning, and clear.

- **offloadGroupEvents**: Used only under the guidance of Oracle Support

- **pmemCacheMode**: Setting for PMEM cache. The value can be writethrough or writeback. The default is writethrough. Available with Oracle Exadata System Software release 19.3.0 or later.

- **powerCount**: Count of power supplies, displayed as working/total

- **powerStatus**: Status of the power. The value can be normal, warning, or critical.

- **pwdExpInDays**: The number of days before a user's password expires. Available with Oracle Exadata System Software release 19.1.0 or later.
- **pwdExpWarnInDays**: The number of days before a user's password expires that a warning message is issued during login attempts. Available with Oracle Exadata System Software release 19.1.0 or later.

- **rackName**: The name of the rack

- **ramCacheMaxSize**: The maximum allowable size of the Cell RAM Cache. Available with Oracle Exadata System Software release 18c (18.1.0) or later.

- **ramCacheMode**: Determines whether the Cell RAM Cache is enabled (on) or disabled (off). Available with Oracle Exadata System Software release 18c (18.1.0) or later.

- **ramCacheSize**: The size of the Cell RAM Cache. Available with Oracle Exadata System Software release 18c (18.1.0) or later.

- **releaseImageStatus**: Indicator for knowing whether imaging is successful

- **releaseTrackingBug**: Patch number for the cell software, such as 30441371

- **releaseVersion**: Release number for the cell software, such as 19.3.1.0.0.191018

- **remotePwdChangeAllowed**: Whether or not a user password can be changed remotely through REST services. Available with Oracle Exadata System Software release 19.1.0 or later.

- **rescuePlan**: A list of commands that you can run after a server rescue to restore settings, such as IORM plans, thresholds, and notifications, to the last known values.

- **rollbackVersion**: The inactive image version that the cell maintains. If `patchmgr rollback` is invoked for the cell, the value displayed by `rollbackVersion` is the software version that will be reinstated.

- **rpmVersion**: The RPM version of the cell

- **rsStatus**: Status of Restart Server

- **securityCert**: The certified identity of the cell. Either CA-certified identity or the default self-certified identity.

- **siteName**: The site name for the cell

- **smtpFrom**: User name that appears in the From: header of the alert notifications

- **smtpFromAddr**: Email address that appears in the From: header of the alert notifications. This email address is not authenticated with the email server.

- **smtpPort**: Email server port used to send alert notifications

- **smtpServer**: Domain name or IP address of the email server used to send alert notifications

- **smtpToAddr**: Address to which email is sent. It can be a comma-delimited list in quotation marks to allow multiple subscribers to alerts.

- **smtpUseSSL**: Specification to use Secure Socket Layer (SSL) encryption for alert notifications.

- **snmpSubscriber**: List of hosts that subscribe to the SNMP alert notifications

- **snmpUser**: Defines the user who receives SNMP alerts

- **status**: Status of the cell
- syslogConf: Designates syslog messages that should be forwarded to a specified management server. Uses the following syntax for the attribute, where selector is the message type, and node is the specified server:

```
syslogconf = (selector @node', [, 'selector @node']... )
```

Both selector and node follow syslog.conf standard syntax rules.

- syslogFormat: A string representing the format for syslog messages. Available with Oracle Exadata System Software release 19.1.0 or later.

- temperatureReading: Current temperature (Celsius) of the cell obtained from the BMC.

- temperatureStatus: Status of the temperature. The value can be normal, warning, or critical.

- traceLevel: The level for which trace messages are written. The default is FINE. The value can be:
  - A valid Java logging level
    * SEVERE
    * WARNING
    * INFO
    * FINE
    * FINER
    * FINEST
  - A valid Oracle Diagnostic Logging (ODL) logging level
    * INCIDENT_ERROR:1
    * ERROR:1
    * WARNING:1
    * NOTIFICATION:1
    * NOTIFICATION:16
    * TRACE:1
    * TRACE:16
    * TRACE:32

- upTime: Time (days, hours:minutes) since the system was restarted

- usbStatus: Status of the USB device

Examples

The following example shows the DESCRIBE command with the CELL object.

**Example 8-92  Describing the CELL Object**

CellCLI> DESCRIBE CELL
<table>
<thead>
<tr>
<th>name</th>
<th>modifiable</th>
</tr>
</thead>
<tbody>
<tr>
<td>accessLevelPerm</td>
<td>modifiable</td>
</tr>
<tr>
<td>accessLevelTemp</td>
<td>modifiable</td>
</tr>
<tr>
<td>accountLockInDays</td>
<td>modifiable</td>
</tr>
<tr>
<td>bbuStatus</td>
<td></td>
</tr>
<tr>
<td>cellsrvStatus</td>
<td></td>
</tr>
<tr>
<td>cellVersion</td>
<td></td>
</tr>
<tr>
<td>comment</td>
<td>modifiable</td>
</tr>
<tr>
<td>cpuCount</td>
<td></td>
</tr>
<tr>
<td>dbPerfDataSuppress</td>
<td>modifiable</td>
</tr>
<tr>
<td>diagHistoryDays</td>
<td>modifiable</td>
</tr>
<tr>
<td>diagPackEmailAttach</td>
<td>modifiable</td>
</tr>
<tr>
<td>diagPackUploadEnabled</td>
<td>modifiable</td>
</tr>
<tr>
<td>doNotServiceLEDStatus</td>
<td></td>
</tr>
<tr>
<td>eighthRack</td>
<td>modifiable</td>
</tr>
<tr>
<td>emailFormat</td>
<td>modifiable</td>
</tr>
<tr>
<td>emailSubscriber</td>
<td>modifiable</td>
</tr>
<tr>
<td>events</td>
<td>modifiable</td>
</tr>
<tr>
<td>exacliEnabled</td>
<td>modifiable</td>
</tr>
<tr>
<td>fanCount</td>
<td></td>
</tr>
<tr>
<td>fanStatus</td>
<td></td>
</tr>
<tr>
<td>flashCacheCompress</td>
<td>modifiable</td>
</tr>
<tr>
<td>flashCacheMode</td>
<td>modifiable</td>
</tr>
<tr>
<td>httpsAccess</td>
<td>modifiable</td>
</tr>
<tr>
<td>id</td>
<td></td>
</tr>
<tr>
<td>interconnect1</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect2</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect3</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect4</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect5</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect6</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect7</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnect8</td>
<td>modifiable</td>
</tr>
<tr>
<td>interconnectCount</td>
<td></td>
</tr>
<tr>
<td>iormBoost</td>
<td></td>
</tr>
<tr>
<td>IOTimeoutThreshold</td>
<td>modifiable</td>
</tr>
<tr>
<td>ipaddress1</td>
<td></td>
</tr>
<tr>
<td>ipaddress2</td>
<td></td>
</tr>
<tr>
<td>ipaddress3</td>
<td></td>
</tr>
<tr>
<td>ipaddress4</td>
<td></td>
</tr>
<tr>
<td>ipaddress5</td>
<td></td>
</tr>
<tr>
<td>ipaddress6</td>
<td></td>
</tr>
<tr>
<td>ipaddress7</td>
<td></td>
</tr>
<tr>
<td>ipaddress8</td>
<td></td>
</tr>
<tr>
<td>kernelVersion</td>
<td></td>
</tr>
<tr>
<td>location</td>
<td>modifiable</td>
</tr>
<tr>
<td>locatorLEDStatus</td>
<td></td>
</tr>
<tr>
<td>makeModel</td>
<td></td>
</tr>
<tr>
<td>memoryGB</td>
<td></td>
</tr>
<tr>
<td>metricCollection</td>
<td>modifiable</td>
</tr>
<tr>
<td>metricHistoryDays</td>
<td>modifiable</td>
</tr>
<tr>
<td>msStatus</td>
<td></td>
</tr>
<tr>
<td>notificationMethod</td>
<td>modifiable</td>
</tr>
<tr>
<td>notificationPolicy</td>
<td>modifiable</td>
</tr>
<tr>
<td>offloadGroupEvents</td>
<td>modifiable</td>
</tr>
</tbody>
</table>
### 8.7.5.5 DESCRIBE CELLDISK

#### Purpose

The `DESCRIBE CELLDISK` command displays a list of attributes for the `CELLDISK` object type.

#### Syntax

```
DESCRIBE CELLDISK
```
Usage Notes

The attributes displayed by the `DESCRIBE CELLDISK` command can include:

- **comment**: User comment for the cell disk.
- **creationTime**: Time stamp when the cell disk was created.
- **deviceName**: Operating system device name of the LUN used by the cell disk.
- **devicePartition**: Operating system device name of the partition that is used by the cell disk.
- **diskType**: The type of disk.
- **errorCount**: Number of errors that occurred on the cell disk.
- **flushError**: Errors reported by while flushing the flash cache.
- **flushStatus**: The current status of the flash cache flush operation.
- **freeSpace**: Amount of unused space available on the cell disk.
- **id**: Global unique identifier (GUID) that is generated when the cell disk is created.
- **physicalDisk**: Name of the physical disk on which the cell disk is located.
- **size**: Total size of the cell disk.
- **status**: Current status of the cell disk, such as `normal` or `importRequired`.

Example 8-93  Describing the CELLDISK Object

This shows the `DESCRIBE` command with the `CELLDISK` object.

```
CellCLI> DESCRIBE CELLDISK  
   name                  modifiable
   comment               modifiable
   creationTime          
   deviceName            
   devicePartition       
   diskType              
   errorCount            
   flushError            hidden
   flushStatus           hidden
   freeSpace             
   id                    
   physicalDisk          
   size                  
   status                
```

Related Topics

- **ALTER CELLDISK**
- **CREATE CELLDISK**
8.7.5.6 DESCRIBE DATABASE

Purpose
Displays the specified attributes for active databases.

Syntax
DESCRIBE DATABASE

Usage Notes
The attributes for the DESCRIBE DATABASE command include the following:

• databaseID: The unique identifier for the database.
• flashCacheLimit: Specifies a "soft" maximum size in flash cache; if the flash cache is not full, a database can exceed the flashCacheLimit value.
• flashCacheMin: The minimum size in the flash cache that is guaranteed for a database even if the blocks are cold.
• flashCacheSize: The size specified for the flash cache.
• iormShare: The IORM database share number.
• lastRequestTime: The time stamp of the last request from the database.
• name: The database name.
• pmemCacheLimit: Specifies a "soft" maximum size in PMEM cache; if the PMEM cache is not full, a database can exceed the pmemCacheLimit value.
• pmemCacheMin: The minimum size in the PMEM cache that is guaranteed for a database even if the blocks are cold.
• pmemCacheSize: The size specified for the PMEM cache.
• profile: The IORM profile associated with the database.

Examples
The following example shows the DESCRIBE command with the DATABASE object.

Example 8-94  Describing the DATABASE Object

```
CellCLI> DESCRIBE DATABASE
    name
    databaseID
    flashCacheLimit
    flashCacheMin
    flashCacheSize
    iormShare
    lastRequestTime
    pmemCacheLimit
    pmemCacheMin
    pmemCacheSize
    profile
```
8.7.5.7 DESCRIBE DISKMAP

**Purpose**
Displays the grid disk attributes for a physical disk.

**Syntax**
DESCRIBE DISKMAP

**Usage Notes**
The attributes displayed by the DESCRIBE DISKMAP command can include:

- **celldisk**: The cell disk name
- **devicePartition**: The disk partition name
- **gridDisks**: The name of the grid disks associated with the disk
- **name**: The disk name.
- **physicalSerial**: The serial number of the disk
- **physicalSize**: The size of the disk
- **slotNumber**: The slot number of the disk
- **status**: The disk status

**Examples**
The following example shows the DESCRIBE command with the DISKMAP object.

**Example 8-95  Describing the DISKMAP Object**

CellCLI> DESCRIBE CELLDISK

name
celldisk
devicePartition
gridDisks
physicalSerial
physicalSize
slotNumber
status

8.7.5.8 DESCRIBE FLASHCACHE

**Purpose**
The DESCRIBE FLASHCACHE command displays a list of attributes for the FLASHCACHE object type.
Syntax

DESCRIBE FLASHCACHE

Usage Notes

The attributes displayed by the DESCRIBE FLASHCACHE command can include:

- `cellDisk`: Cell disk names that contain Exadata Smart Flash Cache.
- `creationTime`: Time stamp when the Exadata Smart Flash Cache was created.
- `degradedCelldisks`: List of cell disks configured for cache but not currently available.
- `effectiveCacheSize`: Usable cache size after deducting space on unavailable cell disks.
- `id`: Global unique identifier (GUID) that is generated when the Exadata Smart Flash Cache is created.
- `name`: Unique name of the Exadata Smart Flash Cache.
- `size`: Total size of the Exadata Smart Flash Cache.
- `status`: Current status of the Exadata Smart Flash Cache, such as normal, warning or critical.

Examples

The following example shows the DESCRIBE command with the FLASHCACHE object.

**Example 8-96 Describing the FLASHCACHE Object**

```plaintext
CellCLI> DESCRIBE FLASHCACHE

<table>
<thead>
<tr>
<th>attribute</th>
<th>modifiable</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
<td></td>
</tr>
<tr>
<td>cellDisk</td>
<td></td>
</tr>
<tr>
<td>creationTime</td>
<td></td>
</tr>
<tr>
<td>degradedCelldisks</td>
<td></td>
</tr>
<tr>
<td>effectiveCacheSize</td>
<td></td>
</tr>
<tr>
<td>id</td>
<td></td>
</tr>
<tr>
<td>size</td>
<td></td>
</tr>
<tr>
<td>status</td>
<td></td>
</tr>
</tbody>
</table>
```

8.7.5.9 DESCRIBE FLASHCACHECONTENT

**Purpose**

The DESCRIBE FLASHCACHECONTENT command displays a list of attributes for the FLASHCACHECONTENT object type.

**Syntax**

DESCRIBE FLASHCACHECONTENT
Usage Notes

The attributes displayed by the DESCRIBE FLASHCACHECONTENT command can include:

- **cachedKeepSize**: Size, in bytes, cached in keep mode for this object.
- **cachedSize**: Size, in bytes, cached for this object.
- **cachedWriteSize**: Size, in bytes, of cached data for this object in write-back flash cache that has not yet been written to hard disk.
- **columnarCacheSize**: Size, in bytes, cached in Hybrid Columnar Compression (HCC) format for this object.
- **columnarKeepSize**: Size, in bytes, cached in Hybrid Columnar Compression (HCC) format that is in keep mode for this object.
- **dbID**: Database unique name identifier.
- **dbUniqueName**: Database unique name.
- **hitCount**: Number of I/Os which read data from flash cache for this object.
- **hoursToExpiration**: Time before this object is downgraded from keep section, if not accessed again.
- **missCount**: Number of I/Os which read data from disk for this object.
- **objectNumber**: Generally, specifies the dictionary object number of the database segment that corresponds to the FLASHCACHECONTENT object.
  
  The value 4294967293 is reserved to indicate that the corresponding FLASHCACHECONTENT object contains cached redo log data.
- **tableSpaceNumber**: Tablespace number for the object number.

Examples

The following example shows the DESCRIBE command with the FLASHCACHECONTENT object.

**Example 8-97  Describing the FLASHCACHECONTENT Object**

```cellcli
CellCLI> DESCRIBE FLASHCACHECONTENT
cachedKeepSize
cachedSize
cachedWriteSize
columnarCacheSize
columnarKeepSize
dbID
dbUniqueName
hitCount
hoursToExpiration
missCount
objectNumber
tableSpaceNumber
```
8.7.5.10 DESCRIBE FLASHLOG

Purpose
The DESCRIBE FLASHLOG command displays a list of attributes for the FLASHLOG object type.

Syntax
DESCRIBE FLASHLOG

Usage Notes
The attributes displayed by the DESCRIBE FLASHLOG command can include:

• cellDisk: Names of the cell disks that contain Oracle Exadata Smart Flash Log.
• creationTime: Timestamp when Oracle Exadata Smart Flash Log was created.
• degradedCelldisks: List of cell disks configured for Oracle Exadata Smart Flash Log, but not currently available.
• effectiveSize: Size of available Oracle Exadata Smart Flash Log after deducting space on unavailable cell disks.
• efficiency: Efficiency of Oracle Exadata Smart Flash Log expressed as a percentage.
• id: Global unique identifier (GUID) generated when Oracle Exadata Smart Flash Log is created.
• name: Unique name of Oracle Exadata Smart Flash Log.
• size: Total size of Oracle Exadata Smart Flash Log.
• status: Current status of Oracle Exadata Smart Flash Log, such as normal, warning or critical. Status normal indicates all flash disks are available. Status warning indicates some flash disks are not available. Status critical indicates all flash disks are unavailable.

Examples
The following example shows the DESCRIBE command with the FLASHLOG object.

Example 8-98 Describing the FLASHLOG Object

CellCLI> DESCRIBE FLASHLOG

name
cellDisk
creationTime
degradedCelldisks
effectiveSize
efficiency
id
size
status
8.7.5.11 DESCRIBE GRIDDISK

Purpose

The **DESCRIBE GRIDDISK** command displays a list of attributes for the **GRIDDISK** object type.

Syntax

```
DESCRIBE GRIDDISK
```

Usage Notes

The attributes for the **DESCRIBE GRIDDISK** command include the following:

- **asmDeactivationOutcome**: Indicator whether a grid disk can be deactivated without loss of data. A value of **YES** indicates the grid disk can be deactivated without data loss.
- **asmDiskgroupName**: Name of the Oracle ASM disk group.
- **asmDiskName**: Name of the Oracle ASM disk.
- **asmDiskRepairTime**: The amount of time the grid disk can remain offline before it is dropped by Oracle ASM.
- **asmDiskSize**: Size of the Oracle ASM disk. This attribute is available in Oracle Exadata System Software release 12.1.2.3.0 and later.
- **asmFailGroupName**: Name of the Oracle ASM failure group.
- **asmModeStatus**: Indicator shows the current Oracle ASM usage of a grid disk. Statuses are **ONLINE**, **OFFLINE**, **DROPPED**, **UNUSED**, **SYNCING**, or **UNKNOWN**.
- **availableTo**: Names of the clients that can access this grid disk.
- **cachedBy**: The name of the flash disks that are currently caching data for this grid disk for write-back flash cache.
- **cachingPolicy**: The flash caching policy for this grid disk. Values are **default** or **none**.
  - **default** means data for this grid disk uses the flash cache.
  - **none** means data for this grid disk do not use flash cache.
  The caching policy can be set when creating a grid disk, or using the **ALTER GRIDDISK** command.
- **cellDisk**: Name of the cell disk that contains the grid disk.
- **comment**: User-supplied text string.
- **creationTime**: Time stamp when the grid disk was created.
- **diskType**: The type of disk.
- **errorCount**: Count of hardware errors detected by the cell disk containing this grid disk.
- **id**: Global unique identifier (GUID) that is generated when the grid disk is created.
• **name**: Unique name of the grid disk.
• **size**: Total size of the grid disk.
• **sparse**: Whether the grid disk is a sparse disk.
• **status**: Current status of the grid disk, such as active, inactive, not present or importRequired.
• **virtualSize**: The size of the disk group for the sparse grid disks.

**Note:**
The `asmDeactivationOutcome` and `asmModeStatus` attributes must be explicitly specified when using the `LIST GRIDDISK` command.

### Example 8-99  Describing the GRIDDISK Object

This example shows the `DESCRIBE` command with the `GRIDDISK` object.

```sh
CellCLI> DESCRIBE GRIDDISK

name                      modifiable
asmDeactivationOutcome    hidden
asmDiskgroupName         hidden
asmDiskName               hidden
asmDiskRepairTime         hidden
asmDiskSize               hidden
asmFailGroupName          hidden
asmModeStatus             hidden
availableTo              modifiable
cachedBy                 modifiable
cachingPolicy            modifiable
cellDisk                 modifiable
comment                   modifiable
creationTime             modifiable
diskType                 modifiable
errorCount               modifiable
id                        modifiable
size                      modifiable
sparse                    modifiable
status                    modifiable
virtualSize               modifiable
```

**Related Topics**

• **CREATE GRIDDISK**
• **ALTER GRIDDISK**
8.7.5.12 DESCRIBE IBPORT

Purpose

The DESCRIBE IBPORT command displays a list of attributes for the IBPORT object type on systems that use InfiniBand Network Fabric.

Note:

This command does not apply to Oracle Exadata Database Machine X8M systems.

Syntax

DESCRIBE IBPORT

Usage Notes

The attributes for the DESCRIBE IBPORT command can include the following:

- **activeSlave**: Indicator whether the port is currently the active port for the bonded IP.
- **dataRate**: The data rate of the InfiniBand Network Fabric port.
- **hcaFWVersion**: The version of the host channel adapter firmware.
- **id**: The Global unique identifier (GUID) of the InfiniBand Network Fabric port.
- **lid**: The local identifier of the InfiniBand Network Fabric port. It is unique within the subnet, and the 16-bit identifiers are used within a network by switches for routing.
- **linkDowned**: The number of times the port training state machine has failed the link error recovery process, and halted the link.
- **linkIntegrityErrs**: The number of link integrity errors.
- **linkRecovers**: The number of times the port training state machine has successfully completed the link error recovery process.
- **name**: The name of the InfiniBand Network Fabric port.
- **physLinkState**: The physical link state.
- **portNumber**: The port number of the InfiniBand Network Fabric port.
- **rcvConstraintErrs**: The number of received constraint errors experienced by the InfiniBand Network Fabric port.
- **rcvData**: The number of 32-bit data words received by the InfiniBand Network Fabric port.
- **rcvErrs**: The number of packets received at the InfiniBand Network Fabric port containing an error.
- **rcvRemotePhysErrs**: The number of physical errors experienced at the InfiniBand Network Fabric port.
• **status**: The link status.

• **symbolErrs**: The number of minor link errors experienced at the InfiniBand Network Fabric port.

• **vl15Dropped**: The number of incoming VL15 packets dropped at the InfiniBand Network Fabric port due to resource limitations, such as lack of buffers.

• **xmtConstraintErrs**: The number of transmitted constraint errors experienced at the InfiniBand Network Fabric port.

• **xmtData**: The number of 32-bit data words transmitted on the InfiniBand Network Fabric port.

• **xmtDiscards**: The number of outbound packets discarded by the InfiniBand Network Fabric port because the port was down or congested.

**Example 8-100 Describing the IBPORT Object**

The following example shows the **DESCRIBE** command with the **IBPORT** object on systems that use InfiniBand Network Fabric.

```
CellCLI> DESCRIBE IBPORT
name
activeSlave
dataRate
hcaFWVersion
id
lid
linkDowned
linkIntegrityErrs
linkRecovers
physLinkState
portNumber
rcvConstraintErrs
rcvData
rcvErrs
rcvRemotePhysErrs
status
symbolErrs
vl15Dropped
xmtConstraintErrs
xmtData
xmtDiscards
```

### 8.7.5.13 DESCRIBE IORMPLAN

**Purpose**

The **DESCRIBE IORMPLAN** command displays a list of attributes for the **IORMPLAN** object type.

**Syntax**

```
DESCRIBE IORMPLAN
```
Usage Notes
The attributes for the DESCRIBE IORMPLAN command can include the following:

- **catPlan**: Allocation plan for the categories set up in the databases using the cell.
- **dbPlan**: Allocation plan for the databases using the cell.
- **name**: Unique name of the interdatabase plan. The name value is automatically set to `cellname_IORMPLAN`.
- **objective**: Optimization mode for IORM.
- **status**: Current status of the interdatabase plan, either active or inactive.

Examples
The following example shows the DESCRIBE command with the IORMPLAN object.

**Example 8-101   Describing the IORMPLAN Object**

```
CellCLI> DESCRIBE IORMPLAN

name               catPlan                 modifiable
                  dbPlan                  modifiable
                  objective               modifiable
                  status
```

Related Topics
- **ALTER IORMPLAN**

8.7.5.14 DESCRIBE KEY

**Purpose**
The DESCRIBE KEY command displays a list of attributes for the KEY object type.

**Syntax**
```
DESCRIBE KEY
```

**Usage Notes**
The attributes for the DESCRIBE KEY command can include the following:

- **key**: Random hexadecimal string used to assign client keys.
- **name**: Name of the key. The value of this field is not displayed with LIST.
- **type**: The type of key.

**Examples**
The following example shows the DESCRIBE command with the KEY object.

```
```
Example 8-102   Describing the KEY Object

CellCLI> DESCRIBE KEY

name key modifiable
     type modifiable

Related Topics

• CREATE KEY

8.7.5.15 DESCRIBE LUN

Purpose

The DESCRIBE LUN command displays a list of attributes for the LUN object type.

Syntax

DESCRIBE LUN

Usage Notes

The attributes for the DESCRIBE LUN command can include the following:

• cellDisk: The name of the flash disk, for example FD_02_rack1celadm10. Not used for hard disks.
• deviceName: Operating system device name for the LUN. For example, /dev/c1d5
• diskType: The type of disk.
• errorCount: Number of errors on this LUN.
• id: Identifier assigned by the system.
• isSystemLun: Indicator whether the disk is a system disk. If value is TRUE, then the disk is a system disk. If the value is FALSE, then the disk is not a system disk, and only has data on it.
• lunSize: Raw size of the LUN before being converted to a cell disk.
• lunUID: Unique identifier assigned by the system.
• lunWriteCacheMode: Status of LUN write cache. The status can be in Write Through Mode or Write Back Mode.
• name: Unique name assigned to the LUN. This might be different (or extended from) the LUN ID if the ID is not unique.
• overProvisioning: Indicator of the percentage of over-provisioned blocks in flash storage that are still available for a particular LUN. This attribute is only used for flash disks.
• physicalDrives: Physical disk names that form the LUN.
• raidLevel: Value of the RAID level that is used on the LUN. For example: RAID 0.
• status: Status of the LUN, which can be normal, warning, or critical.
Examples

The following example shows the `DESCRIBE` command with the `LUN` object.

Example 8-103  Describing the LUN Object

```
CellCLI> DESCRIBE LUN

name
cellDisk
deviceName
diskType
errorCount
id
isSystemLun
lunSize
lunUID
lunWriteCacheMode
overProvisioning
physicalDrives
raidLevel
status
```

Related Topics

•  `LIST DISKMAP`

8.7.5.16 DESCRIBE METRICCURRENT

Purpose

The `DESCRIBE METRICCURRENT` command displays a list of attributes for the `METRICCURRENT` object type.

Syntax

```
DESCRIBE METRICCURRENT
```

Usage Notes

The attributes for the `DESCRIBE METRICCURRENT` command can include the following:

•  `alertState`: Indicator of the alert state. Values are `normal`, `warning`, or `critical`.
•  `collectionTime`: Time stamp when the metric value was collected
•  `metricObjectName`: Name of the object, such as cell disk, grid disk, and consumer group, being measured
•  `metricType`: Specification for how the statistic was created or defined
•  `metricValue`: Value of the metric when it was collected
•  `name`: Unique name of the current metric
•  `objectType`: Type of object being measured. Values are:
Examples

The following example shows the DESCRIBE command with the METRICCURRENT object.

Example 8-104  Describing the METRICCURRENT Object

CellCLI> DESCRIBE METRICCURRENT

name
alertState
collectionTime
metricObjectName
metricType
metricValue
objectType

Related Topics

• DESCRIBE METRICDEFINITION

Exadata Metrics
Exadata metrics are recorded observations of important properties or values relating to the Exadata system software.

8.7.5.17 DESCRIBE METRICDEFINITION

Purpose

The DESCRIBE METRICDEFINITION command displays a list of attributes for the METRICDEFINITION object type.

Syntax

DESCRIBE METRICDEFINITION
Usage Notes

The attributes for the DESCRIBE METRICDEFINITION command can include the following:

- **description**: Description of the metric.
- **metricType**: Indicator of how the statistic was created or defined. The options are as follows:
  - `cumulative`: Cumulative statistics since the metric was created.
  - `instantaneous`: Value at the time that the metric is collected.
  - `rate`: Rates computed by averaging statistics over observation periods.
  - `transition`: Transition metrics are collected at the time their value has changed and typically capture important transitions in hardware status.
- **name**: Unique name of the metric definition. (Details follow this list.)
- **objectType**: Type of object being measured. Values are:
  - `CELL`
  - `CELL_FILESYSTEM`
  - `CELLDISK`
  - `FLASHCACHE`
  - `FLASHLOG`
  - `GRIDDISK`
  - `IBPORT`
  - `IORM_CATEGORY`
  - `IORM_CONSUMER_GROUP`
  - `IORM_DATABASE`
  - `IORM_PLUGGABLE_DATABASE`
  - `HOST_INTERCONNECT`
  - `SMARTIO`
- **persistencePolicy**: Amount of time metric values are stored.
- **unit**: Unit for the metric explicitly, and is related to the metric collected:
  - Number
  - % (percentage)
  - F (Fahrenheit)
  - C (Celsius)
  - IO/sec
  - “IO requests”
  - KB
  - KB/sec
The value of the *name* attribute is a composite of abbreviations. The attribute value starts with an abbreviation of the object type on which the metric is defined:

- **CD**. (cell disk)
- **CG**. (IORM consumer group, database-qualified)
- **CL**. (cell)
- **CT**. (IORM category)
- **DB**. (IORM database-level)
- **FC**. (flash cache)
- **FL**. (flash log)
- **GD**. (grid disk)
- **IORM**
- **N**. (network, IBPORT, HOST_INTERCONNECT)
- **PDB**. (IORM pluggable database)
- **SIO**. (Smart IO)

After the abbreviation of the object type, most of the *name* attributes contain one of the following combinations to identify the operation:

- **IO_BY** (I/O amount)
- **IO_RQ** (number of I/O requests)
- **IO_TM** (I/O latency)
- **IO_WT** (I/O wait time)
- **FC_IO_BY** (Flash cache I/O amount)
- **FC_IO_RQ** (Flash cache I/O requests)
- **FD_IO_BY** (Flash disk I/O amount)
- **FD_IO_RQ** (Flash disk I/O requests)
- **FD_IO_TM** (Flash disk latency)
- **FD_IO_UTIL** (Flash disk utilization percentage)

Next, in the name could be **_R** or **_W** for read or write. Following that in the *name* attribute value there might be **_SM** or **_LG** to identify small or large blocks, respectively.
At the end of the name, there could be _SEC to signify per seconds or _RQ to signify per request.

For consumer group and category metrics, read or write details are omitted.

For example:

- `CD_IO_RQ_R_SM` is the number of requests to read small blocks on a cell disk.
- `GD_IO_TM_W_LG` is the microseconds of I/O latency writing large blocks on a grid disk.

Examples

The following example shows the `DESCRIBE` command with the `METRICDEFINITION` object.

**Example 8-105 Describing the METRICDEFINITION Object**

```
CellCLI> DESCRIBE METRICDEFINITION
```

<table>
<thead>
<tr>
<th>name</th>
<th>description</th>
<th>metricType</th>
<th>objectType</th>
<th>persistencePolicy</th>
<th>unit</th>
</tr>
</thead>
</table>

**Related Topics**

- [LIST METRICDEFINITION](#)
- [Exadata Metrics](#)

Exadata metrics are recorded observations of important properties or values relating to the Exadata system software.

**8.7.5.18 DESCRIBE METRICHISTORY**

**Purpose**

The `DESCRIBE METRICHISTORY` command displays a list of attributes for the `METRICHISTORY` object type.

**Syntax**

```
DESCRIBE METRICHISTORY
```

**Usage Notes**

The attributes for the `DESCRIBE METRICHISTORY` command can include the following:

- `alertState`: Indicator of the alert state. Values are normal, warning, or critical.
- `collectionTime`: Time stamp when the metric value was collected
- `metricObjectName`: Name of the object, such as cell disk, grid disk, and consumer group, being measured
- **metricType**: Specification for how the statistic was created or defined
- **metricValue**: Value of the metric when it was collected
- **metricValueAvg**: Average value of the metric
- **metricValueMax**: Maximum value of the metric
- **metricValueMin**: Minimum value of the metric
- **name**: Name of the current metric
- **objectType**: Type of object being measured. Values are:
  - CELL
  - CELL_FILESYSTEM
  - CELLDISK
  - FLASHCACHE
  - FLASHLOG
  - GRIDDISK
  - IBPORT
  - IORM_CATEGORY
  - IORM_CONSUMER_GROUP
  - IORM_DATABASE
  - IORM_PLUGGABLE_DATABASE
  - HOST_INTERCONNECT
  - SMARTIO

**Examples**

The following example shows the `DESCRIBE` command with the `METRICHISTORY` object.

**Example 8-106   Describing the METRICHISTORY Object**

```
CellCLI> DESCRIBE METRICHISTORY

<table>
<thead>
<tr>
<th>name</th>
<th>alertState</th>
<th>collectionTime</th>
<th>metricObjectName</th>
<th>metricType</th>
<th>metricValue</th>
<th>metricValueAvg</th>
<th>metricValueMax</th>
<th>metricValueMin</th>
<th>objectType</th>
</tr>
</thead>
</table>
```
8.7.5.19 DESCRIBE OFFLOADGROUP

Purpose

The DESCRIBE OFFLOADGROUP command displays a list of attributes for the OFFLOADGROUP object type.

Syntax

DESCRIBE OFFLOADGROUP

Usage Notes

The attributes for the DESCRIBE OFFLOADGROUP command can include the following:

- `autoStart`: Whether the offload server associated with the offload group is dynamically started. Value can be either true or false.
- `comment`: An optional comment
- `creationTime`: The time when the offload group was created
- `id`: An identifier for the offload group
- `isSystemGroup`: Whether the offload group was created by the system software. The value can be either true or false.
- `name`: The name of the offload group
- `package`: modifiable
- `runtimeState`: The current state of the offload group process. The value can be running or stopped.

Example 8-107  Describing the OFFLOADGROUP Object

The following example shows ....

```
CellCLI> DESCRIBE OFFLOADGROUP
        name                     autoStart  comment
        creationTime             id         isSystemGroup
        package                 runtimeState
```

8.7.5.20 DESCRIBE PHYSICALDISK

Purpose

The DESCRIBE PHYSICALDISK command displays a list of attributes for the PHYSICALDISK object type.
Syntax

DESCRIBE PHYSICALDISK

Usage Notes

The attributes for the DESCRIBE PHYSICALDISK command can include the following:

- **ctrlFirmware**: The hard disk controller software version
- **ctrlHwVersion**: The hard disk controller hardware version
- **deviceID**: The ID for the physical disk
- **deviceName**: The name of the physical disk device, for example /dev/sdx
- **diskType**: Type of the disk, whether it is a HardDisk, FlashDisk, or M2Disk.
- **enclosureDeviceId**: Identifier for the hard disk enclosure. This attribute is only applicable to Oracle Exadata System Software on Oracle Exadata Storage Server.
- **errCmdTimeoutCount**: The count of execution of commands related to physical disks that timed out, for example, disk firmware upgrade, listing physical disks, and so on.
- **errHardReadCount**: Total count of read errors on a physical disk
- **errHardWriteCount**: Total count of write errors on a physical disk
- **errorCount**: The sum of all known error counts for a physical disk
- **errOtherCount**: Total error count of all other (unknown) errors for a physical disk
- **errSeekCount**: Total number of disk seek errors
- **flashLifeLeft**: The percentage of flash disk life left for a disk
- **hotPlugCount**: Total number of times a disk has been pulled out and reinserted (hot plugged)
- **lastFailureReason**: The reason for the last physical disk failure
- **luns**: List of LUNs converted from this disk. M.2 disks do not have LUNs.
- **makeModel**: Model description provided by the system
- **name**: Unique name of the physical disk
- **notPresentSince**: Date at which the disk was no longer detected
- **physicalFirmware**: The version of the firmware
- **physicalInsertTime**: Time that the disk was inserted
- **physicalInterface**: Interface type used by the hard disk. For example, SAS
- **physicalPort**: (Only applicable for HP models) The physical disk port value
- **physicalRPM**: The RPM value of a physical hard disk. This attribute is also used to determine the disk type (SATA or SAS).
- **physicalSerial**: System-assigned unique ID
- **physicalSize**: Size of the disk in bytes
- **physicalUseType**: Intended use of the disk, for example, Data Drive
- **sectorRemapCount**: Total number of physical disk sectors that have been remapped because of sector failures

- **slotNumber**: Physical location of disk. This attribute is only applicable to Oracle Exadata System Software on Oracle Exadata Storage Server.

- **status**: Status of the physical disk. Values can be:
  - **failed**: The disk has failed. In earlier releases, this status was called critical.
  - **normal**: The disk is functioning normally
  - **not present**: The disk has been removed
  - **peer failure**: Flash disk failure only
  - **poor performance**: The disk is performing poorly
  - **predictive failure**: The disk is expected to fail
  - **write-through caching**: Flash disk caching only.

**Example 8-108 Describing the PHYSICALDISK Object in Oracle Exadata Storage Server**

```sql
CellCLI> DESCRIBE PHYSICALDISK
           name
        ctrlFirmware
        ctrlHwVersion
        deviceId
        deviceName
        diskType
        enclosureDeviceId
        errCmdTimeoutCount
        errHardReadCount
        errHardWriteCount
        errorCount
        errOtherCount
        errSeekCount
        flashLifeLeft
        hotPlugCount
        lastFailureReason
        luns
        makeModel
        notPresentSince
        physicalFirmware
        physicalInsertTime
        physicalInterface
        physicalPort
        physicalRPM
        physicalSerial
        physicalSize
        physicalUseType
        sectorRemapCount
        slotNumber
        status
```
8.7.5.21 DESCRIBE PLUGGABLEDATABASE

Purpose

The DESCRIBE PLUGGABLEDATABASE command displays a list of attributes for the PLUGGABLEDATABASE object type.

Syntax

DESCRIBE PLUGGABLEDATABASE

Usage Notes

The attributes for the DESCRIBE PLUGGABLEDATABASE command can include the following:

- `asmClusterName`: The Oracle ASM cluster name or alias. Available with Oracle Exadata System Software release 19.1.0 or later.
- `containerName`: The name of the container database (CDB)
- `flashCacheLimit`: The specified limit on the Flash cache for this pluggable database (PDB)
- `flashCacheMin`: The specified minimum size of the Flash cache for this PDB
- `flashCacheSize`: The size of the Flash cache specified for this PDB
- `iormLimit`: The disk I/O utilization limit for the PDB. Available with Oracle Exadata System Software release 19.1.0 or later.
- `iormShare`: The IORM share number for the PDB. Available with Oracle Exadata System Software release 19.1.0 or later.
- `name`: The name of the PDB
- `pdbID`: The ID for the PDB
- `pmemCacheLimit`: The specified limit on the PMEM cache for this pluggable database (PDB)
- `pmemCacheMin`: The specified minimum size of the PMEM cache for this PDB
- `pmemCacheSize`: The size of the PMEM cache specified for this PDB

Example 8-109 Describing the PLUGGABLEDATABASE Object

CellCLI> DESCRIBE PLUGGABLEDATABASE
name
asmClusterName
containerName
flashCacheLimit
flashCacheMin
flashCacheSize
iormLimit
iormShare
pdbID
pmemCacheLimit
8.7.5.22 DESCRIBE PMEMCACHE

**Purpose**

The `DESCRIBE PMEMCACHE` command displays a list of attributes for the PMEMCACHE object type.

**Syntax**

```
DESCRIBE PMEMCACHE
```

**Usage Notes**

The attributes displayed by the `DESCRIBE PMEMCACHE` command can include:

- `cellDisk`: Cell disk names used by the PMEM cache.
- `creationTime`: Time stamp when the PMEM cache was created.
- `degradedCelldisks`: List of cell disks configured for the cache but not currently available.
- `effectiveCacheSize`: Usable PMEM cache size after deducting space on unavailable cell disks.
- `id`: Global unique identifier (GUID) that is generated when the PMEM cache is created.
- `name`: Unique name of the PMEM cache.
- `size`: Total size of the PMEM cache.
- `status`: Current status of the PMEM cache, such as normal, warning or critical.

**Examples**

The following example shows the `DESCRIBE` command with the PMEMCACHE object.

**Example 8-110  Describing the PMEMCACHE Object**

```
CellCLI> DESCRIBE PMEMCACHE

name  cellDisk creationTime degradedCelldisks effectiveCacheSize id size status
```

**Chapter 8**

CellCLI Command Reference

8-126
8.7.5.23 DESCRIBE PMEMLOG

Purpose

The DESCRIBE PMEMLOG command displays a list of attributes for the PMEMLOG object type.

Syntax

DESCRIBE PMEMLOG

Usage Notes

The attributes displayed by the DESCRIBE PMEMLOG command can include:

- **cellDisk**: Names of the cell disks that contain PMEMLOG.
- **creationTime**: Timestamp when PMEMLOG was created.
- **degradedCelldisks**: List of cell disks configured for PMEMLOG that are not currently available.
- **effectiveSize**: Size of available PMEMLOG after deducting space on unavailable cell disks.
- **efficiency**: Efficiency of PMEMLOG expressed as a percentage.
- **id**: Global unique identifier (GUID) generated when PMEMLOG is created.
- **name**: Unique name of PMEMLOG.
- **size**: Total size of the PMEMLOG.
- **status**: Current status of PMEMLOG.
  - normal: All PMEM cell disks are available.
  - warning: Some PMEM cell disks are not available.
  - critical: All PMEM cell disks are unavailable.

Example 8-111  Describing the PMEMLOG Object

This example shows the DESCRIBE command with the PMEMLOG object.

```
CellCLI> DESCRIBE PMEMLOG
  name
cellDisk
creationTime
degradedCelldisks
effectiveSize
efficiency
id
size
status
```
8.7.5.24 DESCRIBE QUARANTINE

Purpose

The DESCRIBE QUARANTINE command displays a list of attributes for the QUARANTINE object type.

Syntax

DESCRIBE QUARANTINE

Usage Notes

The attributes for the DESCRIBE QUARANTINE command can include the following:

- asmClusterId: Identifier of the ASM cluster. This attribute is available in Exadata software 12.2.1.1.0 and later.
- catDBPlan: The name of the category plan
- cellsrvChecksum: Checksum of the CELLSRV binary
- clientPID: The process identifier for the client process which crashed the cell
- comment: Comment for the quarantine
- conDbUniqueID: The container database unique ID for the quarantine
- conDbUniqueName: The container database unique name for the quarantine
- crashReason: Reason for the crash
- creationTime: Quarantine creation time
- dbUniqueID: The database unique ID for the quarantine
- dbUniqueName: The database unique name for the quarantine
- fineGrainControl:
  - fineGrainValue:
- incidentID: The incident identifier of the crash that caused the quarantine creation
- interDBPlan: The name of the interdatabase resource plan
- intraDBPlan: The name of the intradatabase resource plan
- ioBytes: The bytes of quarantined disk region. This is applicable to disk region quarantine only.
- ioGridDisk: The grid disk name for quarantined disk region. This is applicable to disk region quarantine only.
- ioOffset: The I/O offset for quarantined disk region. This is applicable to disk region quarantine only.
- name: Identifier of the quarantine
- objectID:
- planLineID: The SQL Plan Line identifier. This is applicable to SQL Plan quarantine only.
- quarantineMode
- quarantinePlan: This is usually SYSTEM
- quarantineReason: The reason for creation of the quarantine
- quarantineType: The type of quarantine created
- remoteHostName: The host name of the remote host that ran the client process that crashed the cell
- rpmVersion: The RPM version of the cell being used when the cell crashed
- sqlID: The SQLID of the SQL statement that crashed a cell
- sqlPlanHashValue: The SQL Plan hash value. This is applicable to SQL Plan quarantine only.

**Example 8-112 Describing the QUARANTINE Object**

```
CellCLI> DESCRIBE QUARANTINE
name
asmClusterId
catDBPlan
cellsrvChecksum
clientPID
comment                 modifiable
conDbUniqueID
conDbUniqueName
crashReason
creationTime
dbUniqueID
dbUniqueName
fineGrainControl
fineGrainValue
incidentID
interDBPlan
intraDBPlan
ioBytes
ioGridDisk
ioOffset
objectID
planLineID
quarantineMode
quarantinePlan
quarantineReason
quarantineType
remoteHostName
rpmVersion
sqlID
sqlPlanHashValue
```

**8.7.5.25 DESCRIBE ROLE**

**Purpose**

The `DESCRIBE ROLE` command displays a list of attributes for the `ROLE` object type.
Syntax

DESCRIBE ROLE

Usage Notes

The attributes for the DESCRIBE ROLE command can include the following:

- **name**: Unique name of the user assigned the role
- **privileges**: Privileges granted to the role

**Example 8-113 Describing the ROLE Object**

CellCLI> DESCRIBE ROLE
    name
    privileges

---

8.7.5.26 DESCRIBE SOFTWAREHISTORY

**Purpose**

The **DESCRIBE SOFTWAREHISTORY** command displays a list of attributes for the **ALERTHISTORY** object type.

**Syntax**

DESCRIBE SOFTWAREHISTORY

**Usage Notes**

The attributes for the **DESCRIBE SOFTWAREHISTORY** command can include the following:

- **name**: The name of the software update
- **status**: The status of the software update

**Example 8-114 Describing the SOFTWAREHISTORY Object**

CellCLI> DESCRIBE SOFTWAREHISTORY
    name
    status

---

8.7.5.27 DESCRIBE SOFTWAREUPDATE

**Purpose**

The **DESCRIBE SOFTWAREUPDATE** command displays a list of attributes for the **SOFTWAREUPDATE** object type.

**Syntax**

DESCRIBE SOFTWAREUPDATE
Usage Notes

The attributes for the DESCRIBE SOFTWAREUPDATE command can include the following:

- **frequency**: The time period in which this software update is performed automatically. The value can be `none`, `daily`, `weekly`, or `biweekly`. The value `none` is available in Oracle Exadata System Software release 19.1.0 or later.
- **name**: The name of the patch to use in the update, or `unknown`. If the name defaults to `unknown`, then when the software update is performed, the most recent patch is chosen for the upgrade.
- **status**: The status of this software update.
- **store**: The URL for the location of the software update file
- **time**: The specified date and time at which the software update should be performed
- **timeLimitInMinutes**: The number of minutes a cell will spend waiting to update the software before canceling and issuing an alert.

Example 8-115  Describing the SOFTWAREUPDATE Object

```
CellCLI> DESCRIBE SOFTWAREUPDATE
      name           modifiable
      status         modifiable
      store          modifiable
      time           modifiable
      timeLimitInMinutes  modifiable
```

8.7.5.28 DESCRIBE THRESHOLD

**Purpose**

The DESCRIBE THRESHOLD command displays a list of attributes for the THRESHOLD object type.

**Syntax**

DESCRIPTION THRESHOLD

**Usage Notes**

The attributes displayed by the DESCRIBE THRESHOLD command can include:

- **comparison**: Operator for comparing the metric value to the threshold value (>, >=, =, <, <=) to determine whether the value violates the threshold.
- **critical**: Limit beyond which the metric value is considered to be in the critical state for generating alerts
- **name**: Unique name of the threshold
- **observation**: Number of measurements over which the rate metric is averaged before being compared with the threshold value
• occurrences: Number of consecutive violations of the threshold limit by the metric value before the appropriate alert is issued

• warning: Limit beyond which the metric value is considered to be in the warning state for generating alerts

Example 8-116  Describing the THRESHOLD Object

CellCLI> DESCRIBE THRESHOLD

name comparison          modifiable
critical              modifiable
observation           modifiable
occurrences            modifiable
warning                modifiable

8.7.5.29 DESCRIBE USER

Purpose
The DESCRIBE USER command displays a list of attributes for the USER object type.

Syntax

DESCRIBE USER

Usage Notes
The attributes displayed by the DESCRIBE USER command can include:

• name: Unique name of the user

• roles: Roles assigned to the user

Example 8-117  Describing the USER Object

CellCLI> DESCRIBE USER

name
roles

8.7.6 DROP

Purpose
The DROP command removes the named objects from the cell or resets a cell.

Syntax

DROP object_type [object_name [, object_name]...] [options]

Usage Notes

• When multiple objects are the target of a DROP command, there is the possibility of partial success. If an error occurs, then the command is interrupted, and the remaining objects are not dropped.
8.7.6.1 DROP ALERTHISTORY

Purpose

The DROP ALERTHISTORY command removes alerts from the alert history of a cell.

Syntax

DROP ALERTHISTORY {ALL | alert1, alert2, ...}

Usage Notes

• In the command, alertN is the name of the alert to be dropped from the history.
• When dropping stateful alerts, you must drop all members of the alert sequence at the same time. If you do not drop all members, then an error is issued by the system.

Examples

The following example shows the DROP ALERTHISTORY command.

Example 8-118  Dropping a Cell Alert History

CellCLI> DROP ALERTHISTORY 1, 2_1, 2_2

8.7.6.2 DROP CELL

Purpose

The DROP CELL command resets a cell to its original state.

Syntax

DROP CELL [ERASE = value] [FORCE]
Usage Notes

- This command is run from within the cell.
- All cell disks, grid disks, and thresholds are dropped. The interdatabase plan is reset to its default state. All cell attributes are set to default values.
- The `<FORCE>` option is required if the grid disks are configured on any cell disks when `<DROP CELL>` is issued. Otherwise, an error is reported.
- Flash cache compression must be disabled before securely erasing a drive.
- The `<ERASE>` option erases the content on the disk by overwriting the content. The values are as follows:
  - `<1pass>`: One pass, and the content is overwritten with zeros. This value is not available for flash drives.
  - `<3pass>`: Three passes, and the content is overwritten with set data patterns. This option follows the recommendations of NNSA. This value is not available for flash drives.
  - `<7pass>`: Seven passes, and the disk is overwritten with set data patterns. This option follows the recommendations from DOD.
- When dropping all cells using the `<1pass>` or `<3pass>` option, it necessary to drop the flash disks first using the `<7pass>` option, and then drop the cells. The following is an example of the commands:

```plaintext
CellCLI> DROP CELLDISK ALL FLASHDISK ERASE=7pass
CellCLI> DROP CELL ERASE=1pass
```

- As of Oracle Exadata System Software release 19.1.0, if you specify to erase hard disks or flash disks using `<1pass>`, `<3pass>`, or `<7pass>` method on Oracle Exadata Database Machine X5 or later, Oracle Exadata System Software automatically invokes Secure Eraser to erase the disks. Secure Eraser determines whether or not the disks can be erased using the better and faster cryptographic erasure method. If some of the disks are eligible, then the cryptographic erasure method is used to erase those disks, and the originally requested method (1/3/7 pass) is used on the other disks. This feature is not used on system disks. See Table 8-3 for a list of the erasure methods available for each type of device.

The following table shows approximate time needed to securely erase a drive using the supported algorithms. When multiple grid disks or cell disks are dropped with the `<ERASE>` option, the command runs in parallel on all disks and flash drives. However, the recommended method of erasing data from a cell is to use Secure Erase. See Securely Erasing Database Servers and Storage Servers in Oracle Exadata Database Machine Security Guide.

### Table 8-2  Estimated Erasure Times for Devices by Erasure Method

<table>
<thead>
<tr>
<th>Type of Device</th>
<th>1pass</th>
<th>3pass</th>
<th>7pass</th>
<th>Cryptographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>600 GB hard drive</td>
<td>1 hour</td>
<td>3 hours</td>
<td>7 hours</td>
<td>not available</td>
</tr>
<tr>
<td>1.2 TB hard drive</td>
<td>1.67 hours</td>
<td>5 hours</td>
<td>11.67 hours</td>
<td>not available</td>
</tr>
<tr>
<td>2 TB hard drive</td>
<td>5 hours</td>
<td>15 hours</td>
<td>35 hours</td>
<td>not available</td>
</tr>
</tbody>
</table>

See Table 8-3 for a list of the erasure methods available for each type of device.
Table 8-2  (Cont.) Estimated Erasure Times for Devices by Erasure Method

<table>
<thead>
<tr>
<th>Type of Device</th>
<th>1pass</th>
<th>3pass</th>
<th>7pass</th>
<th>Cryptographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 TB hard drive</td>
<td>7 hours</td>
<td>21 hours</td>
<td>49 hours</td>
<td>not available</td>
</tr>
<tr>
<td>4 TB hard drive</td>
<td>8 hours</td>
<td>24 hours</td>
<td>56 hours</td>
<td>not available</td>
</tr>
<tr>
<td>8 TB hard drive</td>
<td>13.17 hours</td>
<td>39.5 hours</td>
<td>92.17 hours</td>
<td>1 min</td>
</tr>
<tr>
<td>10 TB hard drive</td>
<td>14 hours</td>
<td>42 hours</td>
<td>98 hours</td>
<td>1 min</td>
</tr>
<tr>
<td>14 TB hard drive</td>
<td>18 hours</td>
<td>54 hours</td>
<td>126 hours</td>
<td>1 min</td>
</tr>
<tr>
<td>22.875 GB flash drive</td>
<td>not available</td>
<td>not available</td>
<td>21 minutes</td>
<td>not available</td>
</tr>
<tr>
<td>93 GB flash drive</td>
<td>not available</td>
<td>not available</td>
<td>32 minutes</td>
<td>not available</td>
</tr>
<tr>
<td>186 GB flash drive</td>
<td>not available</td>
<td>not available</td>
<td>36 minutes</td>
<td>not available</td>
</tr>
<tr>
<td>1.6 TB flash drive</td>
<td>not available</td>
<td>not available</td>
<td>5.5 hours</td>
<td>1 min</td>
</tr>
<tr>
<td>3.2 TB flash drive</td>
<td>not available</td>
<td>not available</td>
<td>8 hours</td>
<td>1 min</td>
</tr>
<tr>
<td>Persistent Memory (PMEM) device</td>
<td>not available</td>
<td>not available</td>
<td>not available</td>
<td>1 min</td>
</tr>
<tr>
<td>4 GB Internal USB</td>
<td>not available</td>
<td>30 minutes</td>
<td>not available</td>
<td>not available</td>
</tr>
<tr>
<td>8 GB Internal USB</td>
<td>not available</td>
<td>1 hour</td>
<td>not available</td>
<td>not available</td>
</tr>
<tr>
<td>150 GB M.2 device</td>
<td>not available</td>
<td>not available</td>
<td>not available</td>
<td>1 minute</td>
</tr>
<tr>
<td>ILOM</td>
<td>not available</td>
<td>not available</td>
<td>not available</td>
<td>1 minute</td>
</tr>
</tbody>
</table>

Example 8-119  Dropping a Cell

```
CellCLI> DROP CELL FORCE
```

Related Topics

- ALTER CELL

8.7.6.3 DROP CELLDISK

Purpose

The DROP CELLDISK command removes all or the named cell disks from the cell.

This command is necessary if a cell disk fails, or it is replaced by a newer model.

Before dropping the cell disk, you should drop its grid disks and the corresponding Oracle ASM disks from the disk groups. The Oracle ASM disks should be dropped before dropping the grid disks.


## Syntax

```
DROP CELLDISK { ALL [FLASHDISK | HARDDISK] | cdisk_name [, cdisk_name]... } 
[ERASE = value [NOWAIT]] [FORCE]
```

## Usage Notes

- If individual cell disks are specified, then the named cell disks (cdisk_name) are dropped.
- If the **LUN** associated with the **CELLDISK** is flagged as automatically created, then that LUN is deleted along with the cell disk.
- If the **ALL** option is specified, then all the cell disks on the cell are removed.
- The **FLASHDISK** option limits the **DROP CELLDISK** command to cell disks that are flash disks.
- The **HARDDISK** option limits the **DROP CELLDISK** command to cell disks that are hard disks.
- If grid disks are configured on the cell disk when **DROP CELLDISK** is issued, then the **FORCE** option must be used or an error is reported. The **FORCE** option causes any grid disks to be dropped first, and then the cell disk is dropped.
- If the specified cell disk includes flash cache, and that flash cache is in writeback mode, then the cell disk cannot be dropped.
- As of Oracle Exadata System Software release 19.1.0, if you specify to erase hard disks or flash disks using **1pass**, **3pass**, or **7pass** method on Oracle Exadata Database Machine X5 or later, Oracle Exadata System Software automatically invokes Secure Eraser to erase the disks. Secure Eraser determines whether or not the disks can be erased using the better and faster cryptographic erasure method. If some of the disks are eligible, then the cryptographic erasure method is used to erase those disks, and the originally requested method (1/3/7 pass) is used on the other disks. This feature is not used on system disks.
- The **ERASE** option erases the content on the disk by overwriting the content. The values are as follows:
  - **1pass**: One pass, and the content is overwritten with zeros. This option is not applicable for flash drives.
  - **3pass**: Three passes, and the content is overwritten with set data patterns. This option follows the recommendations from the National Nuclear Security Administration (NNSA). This option is not applicable for flash drives. This value is not available for flash drives.
  - **7pass**: Seven passes, and the disk is overwritten with set data patterns. This option follows the recommendations from the United States Department of Defense (DOD).

See **Table 8-2** for the approximate erasure times for each disk and erasure method.
- Use the **NOWAIT** option with the **ERASE** option to run the command asynchronously.
When dropping all cell disks using the 1pass or 3pass option, you must drop the flash disks first using the 7pass option, and then drop the cell disks, for example:

```
CellCLI> DROP CELLDISK ALL FLASHDISK ERASE=7pass
CellCLI> DROP CELLDISK ALL ERASE=1pass
```

The following table gives a summary of the secure erasure methods used for each device type. Hard drives, flash devices, and internal USBs are securely erased in parallel: the time required to erase one device is the same as that required for erasing multiple devices of the same kind.

**Table 8-3  Methods Used to Securely Erase Various Devices**

<table>
<thead>
<tr>
<th>Component</th>
<th>Make or Model</th>
<th>Erasure Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard drive</td>
<td>• 8 TB hard drives on Oracle Exadata Database Machine X5</td>
<td>Cryptographic erase</td>
</tr>
<tr>
<td></td>
<td>• All hard drives on Oracle Exadata Database Machine X6 or later</td>
<td></td>
</tr>
<tr>
<td>Hard drive</td>
<td>All other hard drives</td>
<td>1/3/7-Pass erase</td>
</tr>
<tr>
<td>Flash device</td>
<td>Flash devices on Oracle Exadata Database Machine X5 or later</td>
<td>Cryptographic erase</td>
</tr>
<tr>
<td>Flash device</td>
<td>All other flash devices</td>
<td>7-pass erase</td>
</tr>
<tr>
<td>M.2 device</td>
<td>Oracle Exadata Database Machine X7-2 or later</td>
<td>Cryptographic erase</td>
</tr>
<tr>
<td>Persistent Memory (PMEM) device</td>
<td>Oracle Exadata Database Machine X8M</td>
<td>Cryptographic erase</td>
</tr>
</tbody>
</table>

**Example 8-120  Examples of Dropping a Cell Disk**

```
CellCLI> DROP CELLDISK CD_03_cell01
CellCLI> DROP CELLDISK CD_02_cell06 FORCE
CellCLI> DROP CELLDISK ALL
CellCLI> DROP CELLDISK CD_02_cell09 ERASE=1pass NOWAIT
CellDisk CD_02_cell09 erase is in progress
```

**Related Topics**

- DROP CELL
- Securely Erasing Database Servers and Storage Servers

**8.7.6.4 DROP FLASHCACHE**

**Purpose**

The DROP FLASHCACHE command removes Exadata Smart Flash Cache from a cell.
Syntax
DROP FLASHCACHE

Usage Notes
Before dropping flash cache, the data not synchronized with the grid disk (dirty data) must be flushed from flash cache to the grid disks. Not flushing dirty data may cause data loss.

Examples
The following example shows how to remove Exadata Smart Flash Cache from a cell.

Example 8-121 Removing Exadata Smart Flash Cache

CellCLI> DROP FLASHCACHE

Related Topics
• ALTER FLASHCACHE

8.7.6.5 DROP FLASHLOG

Purpose
The DROP FLASHLOG command removes Oracle Exadata Smart Flash Log from a cell.

Syntax
DROP FLASHLOG [FORCE]

Usage Notes
The DROP FLASHLOG command can be run at runtime, but the command does not complete until all redo data on the flash disk is written to hard disk.

If FORCE is not specified, then the DROP FLASHLOG command fails if there is any saved redo. If FORCE is specified, then all saved redo is purged, and Oracle Exadata Smart Flash Log is removed.

Caution:
If DROP FLASHLOG fails due to the existence of saved redo, then do not use the FORCE option unless you are sure that all saved redo is no longer needed for any databases to perform recovery. Contact Oracle Support Services for additional information.

Examples
The following example shows how to remove Exadata Smart Flash Cache from a cell.
8.7.6.6 DROP GRIDDISK

Purpose

The DROP GRIDDISK command removes the named grid disks from the cell or removes all the grid disks specified by the ALL PREFIX option.

Caution:

Before dropping a grid disk that belongs to an Oracle ASM disk group, ensure that the corresponding disk was dropped from the Oracle ASM disk group.

Syntax

DROP GRIDDISK {ALL [FLASHDISK | HARDDISK ] PREFIX=gdisk_name_prefix , | gdisk_name } [ERASE = value [NOWAIT]] [FORCE]

Usage Notes

• If the gdisk_name is entered, then the name identifies the individual grid disk to be removed. Multiple names can be entered.

• If the ALL PREFIX option is entered, then the gdisk_name_prefix option specifies the prefix assigned to the names of the grid disks to be dropped. The PREFIX option is required when ALL is used.

  Note:
  A comma must be entered after the grid disk prefix when it is followed by the ERASE option.

• The FLASHDISK option limits the DROP GRIDDISK command to grid disks that are flash disks.

• The HARDDISK option limits the DROP GRIDDISK command to grid disks that are hard disks.

• If any of the grid disks are in use when DROP GRIDDISK is issued, then an error is reported. You can use ALTER GRIDDISK with the INACTIVE option to deactivate a grid disk before dropping the grid disk. This action ensures that the grid disk is not in use.

• The FORCE option can be used to force the drop of a grid disk that is in use.

• If the grid disk being dropped was created on a cell disk of type FLASHDISK, then it does not re-create that area or any part of that grid disk or cell disk as FLASHCACHE.
automatically. Use the CREATE FLASHCACHE command to reuse any part of the dropped area for FLASHCACHE.

- The ERASE option erases the content on the disk by overwriting the content. The values are as follows:
  - 1pass: One pass, and the content is overwritten with zeros. This option is not applicable for flash drives. This value is not available for flash drives.
  - 3pass: Three passes, and the content is overwritten with set data patterns. This option follows the recommendations from NNSA. This option is not applicable for flash drives. This value is not available for flash drives.
  - 7pass: Seven passes, and the disk is overwritten with set data patterns. This option follows the recommendations from DOD.

- DROP GRIDDISK ERASE cannot be used for PMEM grid disks.

- When dropping all grid disks using the 1pass or 3pass option, it is necessary to drop the flash disks first using the 7pass option, and then drop the grid disks. The following is an example of the commands:

  CellCLI> DROP GRIDDISK ALL FLASHDISK PREFIX=data, ERASE=7pass
  CellCLI> DROP GRIDDISK ALL PREFIX=data, ERASE=1pass

- Use the NOWAIT option with the ERASE option to run the command asynchronously.

### Example 8-123 Examples of Dropping a Grid Disk

CellCLI> ALTER GRIDDISK data01_CD_03_cell101 INACTIVE

CellCLI> DROP GRIDDISK data01_CD_03_cell101

CellCLI> DROP GRIDDISK ALL PREFIX=data01

CellCLI> DROP GRIDDISK data02_CD_04_cell101 FORCE

CellCLI> DROP GRIDDISK data02_CD_04_cell101 ERASE=1pass

GridDisk data02_CD_04_cell101 successfully dropped

CellCLI> DROP GRIDDISK ALL FLASHDISK PREFIX=DATA, ERASE=7pass

CellCLI> DROP GRIDDISK ALL PREFIX=DATA, ERASE=3pass

**Related Topics**

- DROP CELL
- Dropping a Disk from an Oracle ASM Disk Group
  You can drop a grid disk from a disk group.
- ALTER GRIDDISK

### 8.7.6.7 DROP PMEMCACHE

**Purpose**

The DROP PMEMCACHE command removes the PMEM Cache from a cell.
Syntax

DROP PMEMCACHE

Usage Notes

If the PMEM Cache is in write-back mode, then before dropping the PMEM cache, any data not synchronized with the grid disk (dirty data) must be flushed from PMEM cache to the disks. Failure to flush dirty data can cause data loss.

Example 8-124  Removing PMEM Cache from a Storage Server

DROP PMEMCACHE

8.7.6.8 DROP PMEMLOG

Purpose

The DROP PMEMLOG command removes PMEM log from a cell disk.

Syntax

DROP PMEMLOG [FORCE]

Usage Notes

You can use the DROP PMEMLOG command at run-time, but the command does not complete until all redo data on PMEM is flushed to disk.

If FORCE is not specified, then the DROP PMEMLOG command fails if there is any saved redo. If FORCE is specified, then all data is purged, and PMEM log is removed.

Caution:

If you are unable to drop the PMEM log due to existing redo information in the log, then retry the command first, before using the FORCE option. The FORCE option should only be used in extreme circumstances because it may cause redo log copies to be out of sync, possibly resulting in database redo data corruption. Contact Oracle Support Services for additional information.

Example 8-125  Removing PMEM Log from a Storage Server

This example shows how to remove PMEM log from a storage server.

CellCLI> DROP PMEMLOG
8.7.6.9 DROP QUARANTINE

Purpose
The DROP QUARANTINE command manually drops a quarantine.

Syntax
DROP QUARANTINE { ALL | quarantine1 [, quarantine2]... }

Usage Notes
In general, a quarantine can be removed if the quarantined entity is not expected to cause more problem to CELLSRV. For example, cell offload for problem SQL statements is disabled, or an Oracle Database patch is applied. Refer to the alert message for the quarantine for more details.

When a cell is patched, all quarantines are automatically dropped. It is not necessary to drop them manually.

Examples
The following example shows the DROP QUARANTINE command.

Example 8-126 Dropping Quarantines

CellCLI> DROP QUARANTINE 1

8.7.6.10 DROP ROLE

Purpose
The DROP ROLE command removes user roles from the cell.

Syntax
DROP ROLE { ALL | role_name1 [, role_name2, ...]} [FORCE]

Usage Notes
The FORCE option drops the role when the role has been granted to a user.

Examples
The following example shows how to drop a role.

Example 8-127 Dropping a Role

CellCLI>DROP ROLE gd_monitor

8.7.6.11 DROP SOFTWAREHISTORY

Purpose
The DROP SOFTWAREHISTORY command removes all history or individual update history.
8.7.6.12 DROP THRESHOLD

Purpose
The DROP THRESHOLD command removes all or the specified thresholds from the cell.

Syntax
DROP THRESHOLD { ALL | threshold_name [, threshold_name ...] }

Examples
The following example shows the DROP THRESHOLD command.

Example 8-129 Dropping Thresholds

CellCLI> DROP THRESHOLD ct_io_wt_rq.interactive
CellCLI> DROP THRESHOLD ALL

Related Topics
• DESCRIBE THRESHOLD

8.7.6.13 DROP USER

Purpose
The DROP USER command removes a user from a cell.

Syntax
DROP USER { ALL | user1 [, user2]... }

Examples
The following example shows how to drop a user.

Example 8-130 Dropping a User

CellCLI> DROP USER agarcia
8.7.7 EXIT

Purpose
The EXIT command exits from the CellCLI utility, and returns control to the operating system prompt.

Syntax
EXIT

EXIT has the same functionality as the QUIT command.

8.7.8 EXPORT CELLDISK

Purpose
The EXPORT CELLDISK command prepares all cell disks or a specified cell disk before moving (importing) the cell disk to a different cell.

Syntax
EXPORT CELLDISK { ALL | cdisk_name }

Usage Notes
To move a cell disk from one cell to another, use the EXPORT CELLDISK and IMPORT CELLDISK commands. Usually, all disks are moved to a new cell if the current cell is failing. First, export the cell disk on one cell. Then, import the exported cell disk using the CellCLI utility on the cell where you moved the physical drive that contains the cell disk.

Using cell disk export and import does not preserve the security configuration details associated with ASM-scoped or DB-scoped security. Consequently, after a successful cell disk export and import, you must remove and reconfigure ASM-scoped or DB-scoped security.

When the EXPORT CELLDISK command is run:

- ALL exports all cell disks on the cells that have normal status.
- If the LUN associated with the cell disk is flagged as automatically-created, then that LUN is deleted as part of the export.
- A successfully exported cell disk has the status attribute set to ImportRequired, and the exported cell disk is displayed in the output of the LIST CELLDISK command.
- The following apply when a cell disk is exported (status='ImportRequired') before it is imported:
  - You can change the name and comment attributes.
  - You can drop the cell disk.
  - You cannot create a new grid disk on the cell disk.
- When a disk is exported, any writes from the disk controller cache to the disk are cleared, and the disk is flagged to indicate that the disk was exported. The grid
disks on the disk are no longer visible to Oracle ASM. Any I/Os to the grid disks get errors.

Before exporting a cell disk, the data not synchronized with the grid disk (dirty data) must be flushed from flash cache to the grid disks. Not flushing dirty data may cause data loss.

Examples

The following example shows the **export cell disk** command.

**Example 8-131 Exporting a Cell Disk**

```
CellCLI> EXPORT CELLDISK CD_3_cell01
CellCLI> EXPORT CELLDISK ALL
```

Related Topics

- **ALTER CELLDISK**

8.7.9 GRANT

**Purpose**

The **grant** command sets attributes for privileges and roles.

**Syntax**

```
GRANT object_type [name] TO sub_object_type [sub_object_name]
```

**Usage Notes**

- `object_type` can be as follows:
  - PRIVILEGE
  - ROLE

- The following values can be used for PRIVILEGE object type:
  - `name` is in the following format:
    ```
    { ALL ACTIONS | action } ON { ALL OBJECTS | object }
    [[ ALL ATTRIBUTES | ATTRIBUTES attribute1 [, attribute2, ...] ]] \n    [[ WITH ALL OPTIONS | WITH OPTIONS option1 [, option2, ...] ]] 
    ```
  - The `sub_object_type` must be ROLE.
  - The `sub_object_name` is the name of the role.

- The following can be used for the ROLE object type:
  - `name` is the role name.
  - The `sub_object_type` must be USER.
  - The `sub_object_name` is the name of the user.
8.7.9.1 GRANT PRIVILEGE

Purpose

The `GRANT PRIVILEGE` command sets the access privileges for a role.

Syntax

```
GRANT PRIVILEGE { ALL ACTIONS | action } ON { ALL OBJECTS | object } \ 
{ ALL ATTRIBUTES | ATTRIBUTES attribute1 [, attribute2, ...] } \ 
{ WITH ALL OPTIONS | WITH OPTIONS option1 [, option2, ...] } \ 
TO ROLE { ALL | role1 [, role2, ...] }
```

Usage Notes

- `action` is the command. Examples: `ALTER, CREATE, DESCRIBE, DROP, EXPORT, IMPORT, LIST`.

Notes:
- The `GRANT` and `REVOKE` commands cannot be granted.
- `CREATE USER` and `DROP USER` cannot be granted.
- `CREATE ROLE` and `DROP ROLE` cannot be granted.

- `object` is object type for the action. It can be any CellCLI object. Examples: `CELL, THRESHOLD, PHYSICALDISK, ALERTHISTORY, ROLE`.

- `attribute` are the attributes for the object. To get a list of attributes for an object, run the `LIST object_type` command.

- `option` are the options for the object. Examples: `DETAIL, LIMIT, ORDER BY, WHERE`.

- `role` is the name of the role to grant privileges.

- The `ALL ACTIONS` argument grants privileges for all actions.

- The `ALL OBJECTS` argument grants privileges for all objects.

- The `ALL ATTRIBUTES` argument grants privileges for all attributes.

- The `WITH ALL OPTIONS` argument grants privileges for all options.

- Specifying attributes and `WITH OPTIONS` is optional. If they are not specified, then all attributes and options are granted with the privilege.

Examples

**Example 8-132  Granting Privileges to a Role**

This example shows how to grant privileges to a role.

```
CellCLI> GRANT PRIVILEGE list on alerthistory ATTRIBUTES alertAction,alertMessage \ 
       WITH OPTIONS detail TO ROLE cellmonitor
```
Example 8-133  Granting All Attributes and Options to a Role

This example shows how to grant all attributes and options for a specified action and object to a role.

CellCLI> GRANT PRIVILEGE { ALL ACTIONS | action } ON { ALL OBJECTS | object } to ROLE role1

Example 8-134  Granting All Options with Specified Action, Object and Attributes

This example shows how to grant all options with a specified action, object and attributes to a role.

CellCLI> GRANT PRIVILEGE { ALL ACTIONS | action } ON { ALL OBJECTS | object }  \ ATTRIBUTES attribute1 [, attribute2, ...] to ROLE role1

Example 8-135  Granting All Attributes with Specified Action, Object and Options

This example shows how to grant all attributes with a specified action, object, and options to a role.

CellCLI> GRANT PRIVILEGE { ALL ACTIONS | action } ON { ALL OBJECTS | object }  \ WITH OPTIONS option1 [, option, ...] to ROLE role1

8.7.9.2 GRANT ROLE

Purpose

The GRANT ROLE command sets the role for a user.

Syntax

GRANT ROLE { ALL | role1 [, role2, ...] } TO USER { ALL | user1 [, user2...] }

Usage Notes

- role is the name of the role.
- The ALL argument grants all roles to the user.
- The TO USER ALL argument grants the role to all users.

Examples

Example 8-136  Granting a Role to a User

This example shows how to grant a role to a user.

CellCLI> GRANT ROLE gd_monitor TO USER agarcia
8.7.10 HELP

Purpose

The HELP command displays syntax and usage descriptions for all CellCLI commands.

Syntax

HELP [help_topic]

If no topic argument is provided, HELP displays the name of all available topics. If a topic is specified, then detailed help text is displayed for that topic.

The following example shows examples of the HELP command.

Example 8-137 Display Help Text with the HELP Command

CellCLI> HELP
CellCLI> HELP ALTER
CellCLI> HELP ALTER CELL

8.7.11 IMPORT CELLDISK

Purpose

The IMPORT CELLDISK command reinstates all exported cell disks or an exported cell disk on a cell where you moved the physical drives that contain the cell disks.

The cell disk is typically imported to a different cell than the one from which the cell disk was exported. For example, the physical drive that contains the exported cell disk was moved to a different cell.

When you move a disk with cell disks and grid disks on it from one machine to another, be careful to ensure that the data on it is rebalanced, as per the ASM failure groups. If all disks from one cell are moved to another cell, then there is no need to perform a ASM rebalance, since the entire failure group is moved.

Syntax

IMPORT CELLDISK  { ALL  |  cdisk_name  |  LUN=lun_id  |  cdisk_name  |  LUN=lun_id  }  [,  comment=comment_text]  [FORCE]

Usage Notes

To move a cell disk from one cell to another, use the EXPORT CELLDISK and IMPORT CELLDISK commands. Usually, all disks are moved to a new cell if the current cell is failing. First, export the cell disk on one cell. Then, import the exported cell disk using the CellCLI utility on the cell where you moved the physical drive that contains the cell disk.

Using cell disk export and import does not preserve the security configuration details associated with ASM-scoped or DB-scoped security. Consequently, after a successful cell disk export and import, you must remove and reconfigure ASM-scoped or DB-scoped security.
When the `IMPORT CELLDISK` command is run:

- Either `ALL`, the cell disk name, the LUN ID, or the cell disk name and LUN ID must be specified.
  - `ALL` imports cell disks that have `ImportRequired` status.
  - If the cell disk name is provided and the LUN ID is not provided, then you can import a cell disk by the specified name in cases where Management Server recognizes this cell disk. A recognized cell disk is displayed in the output of `LIST CELLDISK` with status equal to `ImportRequired`.
  - If the LUN ID is provided and the cell disk name is not provided, then the LUN is scanned, and the cell disk is imported. This variation of the command can be used to import a newly-inserted cell disk that was not recognized by Management Server and Cell Server.
  - If the LUN ID and cell disk name are both provided, then the LUN ID is used to import the cell disk, and the name is used to rename the imported cell disk.
- A new value can be entered for the `comment` attribute to update the existing cell disk comment.
- The cell disk name is verified to ensure that the name is unique within the cell. Cell disks can be renamed before import to ensure uniqueness.
- The grid disk names within a cell must be unique. If a physical disk is moved from one cell (`cell_A`) to another cell (`cell_B`) using the `EXPORT` and `IMPORT` commands, then there is a chance that the target cell (`cell_B`) could have two grid disks with identical names. In this case, the cell software automatically resolves the naming conflict by adding a temporary suffix (`_duplicate_name`, `_duplicate_name2`, `_duplicate_name3`, and so on) to the name of one of the grid disks. This additional suffix enables you to refer to a grid disk unambiguously in the CellCLI commands.

It is recommended that you rename a duplicate grid disk on a cell (`cell_B`) with a new permanent unique name using the following command:

```
ALTER GRIDDISK gdname_duplicate_name NAME=new_unique_name
```

If you return the physical disk to the original cell (`cell_A`) or move the disk to another cell rather than renaming the disk, then the grid disk displays its original name.

- The `LIST CELLDISK` command shows which cell disks need to be imported. The command displays output similar to the following:

```
CellCLI> list celldisk
CD_01_cell100     normal
CD_01_cell101     normal
CD_01_cell102     importRequired
CD_01_cell103     importForceRequired
CD_01_cell104     importRequired
CD_01_flash0      normal
CD_01_flash1      normal
CD_01_flash2      normal
CD_01_log00       normal
CD_01_log01       normal
```
• If the cell disk was not successfully exported and moved between cells, then the \texttt{FORCE} option must be specified with \texttt{IMPORT} or an error occurs. Oracle recommends contacting Oracle Support Services before using the \texttt{FORCE} option.

• The \texttt{IMPORT} command checks the disk to determine if it was exported. If it was exported, then the \texttt{IMPORT} command makes the grid disk visible to Oracle ASM. If the disk was not exported, then the \texttt{FORCE} option should be used with the \texttt{IMPORT} command to reconstruct the grid disks on the disk, and make them visible to Oracle ASM.

**Example 8-138  Importing a Cell Disk**
This example shows the \texttt{IMPORT CELLDISK} command. The LUN ID is provided with the \texttt{IMPORT} command to identify the cell disk, and the cell disk name is used to rename the cell disk on the cell where it was imported.

```
CellCLI> IMPORT CELLDISK CD_7_cell04 lun=3
CellCLI> IMPORT CELLDISK ALL
```

**8.7.12 LIST**

**Purpose**

The \texttt{LIST} command displays attributes for Oracle Exadata System Software objects. Objects displayed are identified by name or by filters. The attributes displayed for each object are determined by the specified attribute list.

**Syntax**

```
LIST object_type \[ name | attribute_filters \] [attribute_list] \[DETAIL\] \[ORDER BY attribute [ASC| DESC],[ attribute [ASC| DESC], ...] \[LIMIT integer]\n```

**Usage Notes**

• Using \texttt{LIST} with only an \texttt{object_type} (without the \texttt{DETAIL} option or an attribute list) displays the names of the existing objects of this type and a default list of attributes.
  – For an object type that has a \texttt{status} attribute, the object name and the status are displayed.
  – For the \texttt{METRICHISTORY} object type, the collection time, the object name, and value are displayed.
  – For the \texttt{PHYSICALDISK} and \texttt{LUN} object types, the ID attribute is displayed.
  – For the \texttt{ALERTHISTORY} object type, the time and alert message are displayed.
  – For the \texttt{KEY} object type, the key value is displayed.

• The attributes displayed for each object are determined by the specified attribute list. Attribute values that are strings with embedded blank spaces or tabs must be enclosed in quotation marks.
Attribute filters determine the specific objects that are displayed. Because of the amount of metrics, you should use filters when using the LIST METRICCURRENT or LIST METRICHISTORY commands to narrow the output of the command.

In the default format without the DETAIL option, each object is displayed on a separate line, with successive attribute values separated by tabs in the order of the specified list of attributes.

In the DETAIL format, each attribute of a specific object is displayed on a separate line, with an attribute name followed by its value. If no attribute list is provided, then all attributes that have values are displayed. Blank lines separate each object in the display. DETAIL is similar to the ATTRIBUTES ALL option, only the format is different.

Attributes that are not set are not listed with the DETAIL option. However, attributes that are set to an empty value are listed with the DETAIL option.

The ORDER BY option orders attributes in ascending or descending order. The default is ASC.

The LIMIT option sets a limit on the number of displayed attributes. The maximum value is 100 when LIMIT is used with the ORDER BY option.

- LIST ACTIVEREQUEST
- LIST ALERTDEFINITION
- LIST ALERTHISTORY
- LIST CELL
- LIST CELLDISK
- LIST DATABASE
- LIST DIAGPACK
- LIST DISKMAP
- LIST FLASHCACHE
- LIST FLASHCACHECONTENT
- LIST FLASHLOG
- LIST GRIDDISK
- LIST IBPORT
- LIST IORMPLAN
- LIST IORMPROFILE
- LIST KEY
- LIST LUN
- LIST METRICCURRENT
- LIST METRICDEFINITION
- LIST METRICHISTORY
- LIST OFFLOADGROUP
- LIST OFFLOADPACKAGE
- LIST PHYSICALDISK
- LIST PLUGGABLEDATABASE
- LIST PMEMCACHE
- LIST PMEMLOG
  The LIST PMEMLOG command displays the specified attributes for the PMEMLOG.
- LIST QUARANTINE
- LIST ROLE
- LIST SOFTWAREHISTORY
- LIST SOFTWAREUPDATE
- LIST THRESHOLD
- LIST USER

Related Topics
- Attribute Lists in LIST Command
  You can specify which attributes to display for the LIST command with the optional
  ATTRIBUTES clause.
- Attribute Filters in LIST and ALTER GRIDDISK Commands
  You can specify which objects to display for the LIST and ALTER GRIDDISK
  commands with the optional attribute_filters clause.

### 8.7.12.1 LIST ACTIVEREQUEST

**Purpose**

The LIST ACTIVEREQUEST command displays specified attributes for the outstanding active requests for the cell.

**Syntax**

```
LIST ACTIVEREQUEST  [ name | attribute_filters ] [attribute_list] [DETAIL]
```

**Usage Notes**

The list of attributes that can be displayed is shown in the following example.

**Example 8-139  Listing ACTIVEREQUEST Attributes**

This example shows the LIST command with the ACTIVEREQUEST object.

```
CellCLI> LIST ACTIVEREQUEST 5 DETAIL
```

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
<td>5</td>
</tr>
<tr>
<td>ID</td>
<td>5</td>
</tr>
<tr>
<td>ParentID</td>
<td>5</td>
</tr>
<tr>
<td>dbName</td>
<td>&quot;test DB&quot;</td>
</tr>
<tr>
<td>InstNum</td>
<td>5</td>
</tr>
<tr>
<td>ConsumerGrp</td>
<td>&quot;test group&quot;</td>
</tr>
<tr>
<td>SessID</td>
<td>5</td>
</tr>
<tr>
<td>SerialNum</td>
<td>5</td>
</tr>
<tr>
<td>AsmFileNum</td>
<td>5</td>
</tr>
<tr>
<td>AsmDGNum</td>
<td>5</td>
</tr>
<tr>
<td>FileIncNum</td>
<td>5</td>
</tr>
</tbody>
</table>
8.7.12.2 LIST ALERTDEFINITION

**Purpose**

The `LIST ALERTDEFINITION` command displays all available sources of the alerts on the cell.

**Syntax**

```plaintext
LIST ALERTDEFINITION [ name | attribute_filters ] [attribute_list] [DETAIL]
```

**Usage Notes**

The list of attributes that can be displayed is shown in the following example.

**Example 8-140  Listing ALERTDEFINITION Attributes**

This example shows the `LIST` command with the `ALERTDEFINITION` object.

```
CellCLI> LIST ALERTDEFINITION StatefulAlert.CG_IO_RQ_LG DETAIL
```

```
name:                     StatefulAlert.CG_IO_RQ_LG
alertShortName:           CG_IO_RQ_LG
alertSource:              Metric
alertType:                Stateful
description:              "Threshold Alert"
metricName:               CG_IO_RQ_LG
```

8.7.12.3 LIST ALERTHISTORY

**Purpose**

The `LIST ALERTHISTORY` command displays all alerts that occurred on the cell.

**Syntax**

```plaintext
LIST ALERTHISTORY [ name | attribute_filters ] [attribute_list] [DETAIL]
```

**Usage Notes**

The list of attributes that can be displayed is shown in Example 8-91.
A WHERE clause can include the ageInMinutes attribute to specify the list is limited to those alerts which have the specified age. For example, the following command would show the alerts created in the previous 15 minutes:

CellCLI> LIST ALERTHISTORY WHERE ageInMinutes < 15

The alerthistory attribute can be used to check for stateful and stateless alerts.

Examples

Example 8-141 shows the LIST command with the ALERTHISTORY object.

Example 8-142 shows open stateful and stateless alerts.

Example 8-143 shows only open stateful alerts. The closed stateful alerts are filtered out.

Example 8-144 shows alerts that have not been cleared.

Example 8-141  Listing ALERTHISTORY Attributes

CellCLI> LIST ALERTHISTORY 1671443714 DETAIL
name: 1671443714
alertSequenceID: 1671443714
sequenceBeginTime: 1179185707672
beginTime: "Sat May 18 10:14:16 PDT 2009"
endTime: "Sat May 25 10:14:16 PDT 2009"
severity: critical
alertMessage: "Errors in file svtrc_2840_10.trc (incident=13):"

alertShortName: ADR
alertNotified: 0
examinedBy: johndoe
alertType: stateless

CellCLI> LIST ALERTHISTORY WHERE begintime > 'Jun 1, 2009 11:37:00 AM PDT'

39 2009-10-02T12:26:53-07:00 "ORA-07445: exception encountered: core dump [__kernel l_vsyscall()+5] [6] [0x408C] [[]] [[]]"
40 2009-10-06T23:28:06-07:00 "RS-7445 [unknown_function]
(signum: 6) [[]] [[]] [[]] [[]]"
41 2009-10-07T00:50:42-07:00 "RS-7445 [It will be restarted]
[unknown_function]
(signum: 6) [[]] [[]] [[]] [[]]
42 2009-10-07T02:19:19-07:00 "RS-7445 [unknown_function]
(signum: 6) [[]] [[]] [[]] [[]]

CellCLI> LIST ALERTHISTORY 7 DETAIL
name: 7
alertMessage: "Flash cache mode is set to WriteBack because there is dirty data in the flash"
Example 8-142  Listing Open Stateful and Stateless Alerts

CellCLI> LIST ALERTHISTORY ATTRIBUTES alertsequenceid,name,alerttype \  
              WHERE endtime=null

<table>
<thead>
<tr>
<th>alertsequenceid</th>
<th>name</th>
<th>alerttype</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Stateless</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Stateless</td>
</tr>
<tr>
<td>11</td>
<td>11_1</td>
<td>Stateful</td>
</tr>
</tbody>
</table>

Example 8-143  Listing Open Stateful Alerts

CellCLI> LIST ALERTHISTORY WHERE endtime=null AND alerttype=stateful

Example 8-144  Listing Alerts That Have Not Cleared

CellCLI> LIST ALERTHISTORY WHERE endtime=null

<table>
<thead>
<tr>
<th>alertsequenceid</th>
<th>beginTime</th>
<th>severity</th>
<th>alertAction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2014-11-11T11:08:15-08:00 info</td>
<td>Factory defaults restored for Adapter 0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2014-11-11T11:27:06-08:00 critical</td>
<td>RS-700 [No IP found in Exadata config file] [Check cellinit.ora]</td>
<td></td>
</tr>
<tr>
<td>11_1</td>
<td>2014-12-19T12:01:06-08:00 critical</td>
<td>The HDD disk controller battery has failed. All disk drives have been placed in WriteThrough caching mode. Disk write performance may be reduced. The flash drives are not affected. Battery Serial Number : 1142 Battery Type : ibbu08 Battery Temperature : 39 C Full Charge Capacity : 773 mAh Relative Charge : 83%</td>
<td></td>
</tr>
</tbody>
</table>
Ambient
Temperature : 32 C°

8.7.12.4 LIST CELL

Purpose
The LIST CELL command displays specified attributes of the cell.

Syntax
LIST CELL [ATTRIBUTES attribute_list] [DETAIL]

Usage Notes
• The list of attributes that can be displayed is shown in Example 8-92.
• LIST CELL only lists the local cell. Name and filter options on LIST CELL are not required.
• To monitor the status of cell components, use the LIST command to verify the value of status, fanStatus, temperatureStatus, and powerStatus.

Examples
Example 8-145 shows the LIST command with the CELL object, and the corresponding output.
Example 8-146 shows how to display the status of cell components.
Example 8-147 shows how to display the values of the snmpSubscriber attribute.
Example 8-148 shows how to display the value of the emailFormat attribute.
Example 8-149 shows how to display the value of the locateLEDStatus attribute.
Example 8-150 shows how to display the value of the doNotServiceLEDStatus attribute.
Example 8-151 shows how to display the value of the bbuLearnCycleTime attribute.
Example 8-152 shows how to display the value of the rescuePlan attribute.
Example 8-153 shows how to retrieve the value of the httpsAccess attribute.

Example 8-145  Listing Cell Information

CellCLI> LIST CELL
          cell01        online

Example 8-146  Displaying the Status of Cell Components

CellCLI> LIST CELL ATTRIBUTES name, status, location, -fanStatus, temperatureStatus, powerStatus
          cell01        online     rack5:shelf1     normal   normal   normal
Example 8-147  Displaying the snmpSubscriber Attribute

```bash
CellCLI> LIST CELL ATTRIBUTES snmpSubscriber
((host=server1.example.com,port=3873,community=public, type=asr))
```

Example 8-148  Displaying E-mail Format

```bash
CellCLI> LIST CELL ATTRIBUTES emailFormat
html
```

Example 8-149  Displaying locateLEDStatus

```bash
CellCLI> LIST CELL ATTRIBUTES locateLEDStatus
off
```

Example 8-150  Displaying doNotServiceLEDStatus

```bash
CellCLI> LIST CELL ATTRIBUTES doNotServiceLEDStatus
on
```

Example 8-151  Listing the bbuLearnCycleTime Attribute

```bash
CellCLI> LIST CELL ATTRIBUTES bbuLearnCycleTime
```

Example 8-152  Displaying rescuePlan

```bash
CellCLI> LIST CELL ATTRIBUTES rescuePlan
```

```
CREATE ROLE "admin"
GRANT PRIVILEGE all actions ON diagpack all attributes WITH all options TO ROLE "admin"

CREATE ROLE "diagRole"
GRANT PRIVILEGE download ON diagpack all attributes WITH all options TO ROLE "diagRole"

GRANT PRIVILEGE create ON diagpack all attributes WITH all options TO ROLE "diagRole"

GRANT PRIVILEGE list ON diagpack all attributes WITH all options TO ROLE "diagRole"

ALTER CELL accessLevelPerm="remoteLoginEnabled", diagHistoryDays="7", metricHistoryDays="7", notificationMethod="mail,snmp", notificationPolicy="warning,critical,clear", snmpSubscriber=((host="localhost", port=162, community="public", type=asr)), bbuLearnCycleTime="2016-10-17T02:00:00-07:00", bbuLearnSchedule="MONTH 1 DATE 17 HOUR 2 MINUTE 0", alertSummaryStartTime="2016-09-21T17:00:00-07:00",
```
alertSummaryInterval=weekly, hardDiskScrubInterval=biweekly, hardDiskScrubFollowupIntervalInDays="14"

ALTER IORMPLAN objective=basic

Example 8-153  Displaying the HTTPs Access Control List

This example shows how to view the HTTPs access control list for the Exadata RESTful service.

CellCLI> LIST CELL ATTRIBUTES httpsAccess ALL

The value of ALL is the default value and allows access to all hosts.

8.7.12.5 LIST CELLDISK

Purpose

The LIST CELLDISK command displays attributes for cell disks determined by the specified attributes and filters.

Syntax

LIST CELLDISK [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Describing the CELLDISK Object.

Examples

The following example shows the LIST command with the CELLDISK object, and the corresponding output.

Example 8-154  Listing Cell Disk Attributes

CellCLI> LIST CELLDISK CD_01_cell105 ATTRIBUTES size

557.859375G

CellCLI> LIST CELLDISK WHERE status!=normal ATTRIBUTES name

CD_01_cell103

CellCLI> LIST CELLDISK WHERE DEVICENAME LIKE '/dev/c0d[2-5]'

ATTRIBUTES name, size

CD_01_cell105         557.859375G

CellCLI> LIST CELLDISK CD_01_cell105 DETAIL
8.7.12.6 LIST DATABASE

Purpose
Displays the specified attributes for active databases.

Syntax

LIST DATABASE [name | attribute_filters] [attribute_list] [DETAIL]

Usage Notes

The filters option is an expression that determines which active databases should be listed by the command.

The list of attributes that can be displayed is shown in Example 8-94.

Examples

The following example shows the LIST command with the DATABASE object, and the corresponding output.

Example 8-155 Listing Database Attributes

CellCLI> LIST DATABASE
   DB01

CellCLI> LIST DATABASE DETAIL
   name: DB01
   asmClusterName: SALESDBS_ASMCLUSTER
   databaseID: 1234567656
   lastRequestTime: 2016-10-27T07:46:36-07:00
   profile: GOLD
   flashCacheMin: 4.00390625G
   flashCacheLimit: 4.19921875G
   flashCacheSize: 0
   pmemCacheMin: 2.001953125G
pmemCacheLimit: 2.099609375G
pmemCacheSize: 0

CellCLI> LIST DATABASE ATTRIBUTES NAME, PROFILE
ASM
TEST50 GOLD
TEST100 GOLD
TEST150 SILVER
TEST20 GOLD
TEST200 BRONZE
TEST180 SILVER
TEST175 SILVER
TEST225 BRONZE
TEST230 BRONZE
TEST300
TEST280
TEST245 BRONZE

CellCLI> LIST DATABASE ATTRIBUTES NAME, DATABASEID WHERE PROFILE = 'GOLD'

8.7.12.7 LIST DIAGPACK

Purpose

The LIST DIAGPACK command lists the diagnostic packages in your system, along with their status.

Syntax

LIST DIAGPACK [DETAIL]

Usage Notes

The location of the diagnostic packages is $LOG_HOME.

Examples

Example 8-156  Output of the "list diagpack" Command

This example shows the output of the LIST DIAGPACK command.

CellCLI> LIST DIAGPACK
  scab01cel04_diag_2015_09_30T13_29_06_1.tar.bz2  (7 minutes ago)
  scab01cel04_2015_09_30T13_13_00_2_1.tar.bz2    (23 minutes ago for alert: 2_1)
  scab01cel04_2015_09_30T13_07_10_1_1.tar.bz2    (28 minutes ago for alert: 1_1)

Example 8-157  Output of the "list diagpack" command with the DETAIL option

This example shows the output of the LIST DIAGPACK command with the DETAIL option.
8.7.12.8 LIST DISKMAP

Purpose
Displays the specified grid disk attributes for a physical disk.

Syntax
LIST DISKMAP

Usage Notes
The list of attributes that can be displayed is shown in Example 8-95.

Examples
The following example shows the LIST command with the DISKMAP object, and the corresponding output.

Example 8-158  Listing Grid Disk Attributes for a Physical Disk

CELLCLI> LIST DISKMAP

Name   PhysicalSerial  SlotNumber    Status  PhysicalSize  CellDisk  DevicePartition  GridDisks
27:0   E0XH34          0             normal   559G         CD_00_sgrcel2   /dev/sda3     "DATA_CD_00_sgrcel2, RECO_CD_00_sgrcel2"
27:1   E0XH2S          1             normal   559G         CD_01_sgrcel2   /dev/sdb3     "DATA_CD_01_sgrcel2, RECO_CD_01_sgrcel2"
27:2   E0Z0CS           2            normal   559G         CD_02_sgrcel2   /dev/sdc     "DATA_CD_02_sgrcel2, DBFS_CD_02_sgrcel2, RECO_CD_02_sgrcel2"
8.7.12.9 LIST FLASHCACHE

Purpose
The LIST FLASHCACHE command displays attributes for the Exadata Smart Flash Cache determined by the specified attributes.

Syntax
LIST FLASHCACHE [attribute_list] [DETAIL]

Usage Notes
The list of attributes that can be displayed is shown in Example 8-96.

Examples
The following example shows the LIST command with the FLASHCACHE object, and the corresponding output.

Example 8-159 Listing Exadata Smart Flash Cache Attributes

CellCLI> LIST FLASHCACHE

    raw_FLASHCACHE normal

CellCLI> LIST FLASHCACHE DETAIL

  name:                   raw_FLASHCACHE
  cellDisk:               c9FLASH0,FD_FLASH1_raw,FD_FLASH2_raw
  creationTime:           2012-08-04T15:42:42-07:00
  degradedCelldisks:     
  effectiveCacheSize:    192M
  id:                     8a0adc84-9088-4c4e-8e1c-b6bcbd5cb1ba
  size:                   192M
  status:                 normal

8.7.12.10 LIST FLASHCACHECONTENT

Purpose
The LIST FLASHCACHECONTENT command displays attributes for the Exadata Smart Flash Cache entries determined by the specified attributes.

Syntax
LIST FLASHCACHECONTENT [attribute_filters] [attribute_list] [DETAIL]

Usage Notes
The list of attributes that can be displayed is shown in Example 8-97.
Examples

Example 8-160  Listing Exadata Smart Flash Cache Content Attributes

This example shows the `LIST` command with the `FLASHCACHECONTENT` object, and the corresponding output.

```
CellCLI> LIST FLASHCACHECONTENT DETAIL

cachedKeepSize:             8192
cachedSize:                 16384
cachedWriteSize:            16384
clusterName:                CLUSTER-C1
columnarCacheSize:          0
columnarKeepSize:           0
dbID:                       3557170052
dbNameUnique:               TEST1
hitCount:                   4
missCount:                  
objectNumber:               23102
tableNameNumber:            1

cachedKeepSize:             0
cachedSize:                 983040
cachedWriteSize:            983040
clusterName:                CLUSTER-C1
columnarCacheSize:          0
columnarKeepSize:           0
dbID:                       4325252357
dbUniqueName:               MYODB
hitCount:                   1
missCount:                  1
objectNumber:               4294967295
tableNameNumber:            4
```

Example 8-161  Listing Exadata Smart Flash Cache Content by Database Object

This example shows a database query for an object in a partitioned table, and then lists the flash cache for the same object. In the example, a partitioned table is created in the database, and then queried for the data object numbers of the partitions. The flash cache on Oracle Exadata Storage Server is then queried.

```sql
CREATE TABLE parttabl (c1 number) PARTITION BY RANGE(c1)
(
    PARTITION partt1 VALUES LESS THAN (100),
    PARTITION partt2 VALUES LESS THAN (200)
);

SQL> SELECT SUBSTR(OBJECT_NAME, 0 , 10) OBJ_NAME, SUBOBJECT_NAME,
DATA_OBJECT_ID
            FROM user_objects WHERE OBJECT_NAME LIKE ('PARTT%');

OBJ_NAME   SUBOBJECT_NAME                 DATA_OBJECT_ID
---------- ------------------------------ --------------
PARTTABL PARTT1 63197
PARTTABL PARTT2 63198

CellCLI> LIST FLASHCACHECONTENT WHERE objectNumber=63197 DETAIL
  cachedKeepSize:             0
  cachedSize:                 983040
  cachedWriteSize:            983040
  clusterName:                CLUSTER-C1
  columnarCacheSize:          0
  columnarKeepSize:           0
  dbID:                       3557170052
  dbUniqueName:               TEST1
  hitCount:                   1
  missCount:                  1
  objectNumber:               63197
  tableSpaceNumber:           4

CellCLI> LIST FLASHCACHECONTENT WHERE objectNumber=63198 DETAIL
  cachedKeepSize:             0
  cachedSize:                 16384
  cachedWriteSize:            16384
  clusterName:                CLUSTER-C1
  columnarCacheSize:          0
  columnarKeepSize:           0
  dbID:                       3557170052
  dbUniqueName:               TEST1
  hitCount:                   0
  missCount:                  2
  objectNumber:               63198
  tableSpaceNumber:           4

Example 8-162  Listing Exadata Smart Flash Cache Content for an
ASMCLUSTER Client

Starting with Oracle Exadata System Software release 19.1.0, the dbUniqueName
attribute is qualified with the ASMCLUSTER client name if ASM-scoped security is
configured. This example shows the partial output of the LIST FLASHCACHECONTENT
command for database instances associated with Oracle ASM clusters.

CellCLI> LIST FLASHCACHECONTENT WHERE dbuniquename LIKE 'ASM.*'
  cachedKeepSize:         0
  cachedSize:             65536
  cachedWriteSize:        65536
  columnarCacheSize:      0
  columnarKeepSize:       0
  dbID:                   3334479949
  dbUniqueName:           ASM1.DB1.CDB$ROOT
  hitCount:               0
  missCount:              0
  objectNumber:           75307
  tableSpaceNumber:       1
Example 8-163  Listing Exadata Smart Flash Cache Content for Cached Redo Logs

Starting with Oracle Exadata System Software release 20.1.0, Exadata Smart Flash Cache can contain cached redo logs. Cached redo log entries are identified by objectNumber = 4294967293. This example shows a LIST FLASHCACHECONTENT command that displays details for cached redo logs.

```
CellCLI> LIST FLASHCACHECONTENT -
         ATTRIBUTES dbUniqueName, dbID, tableSpaceNumber, objectNumber, cachedSize -
         WHERE objectNumber = 4294967293

   dbUniqueName: ACME1             3557170052      0        4294967293
   dbUniqueName: MYODB             4325252357      0        4294967293
```

8.7.12.11 LIST FLASHLOG

Purpose

The LIST FLASHLOG command displays attributes for the Oracle Exadata Smart Flash Log entries determined by the specified attributes.
Syntax

LIST FLASHLOG

Usage Notes

The list of attributes that can be displayed is shown in Example 8-98.

Examples

The following example shows the LIST command with the FLASHLOG object, and the corresponding output.

Example 8-164 Listing Oracle Exadata Smart Flash Log Attributes

CellCLI> LIST FLASHLOG

    raw_FLASHLOG normal

CellCLI> LIST FLASHLOG DETAIL

<table>
<thead>
<tr>
<th>name</th>
<th>raw_FLASHLOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>cellDisk</td>
<td>c9FLASH0,FD_FLASH1_raw,FD_FLASH2_raw</td>
</tr>
<tr>
<td>creationTime</td>
<td>2011-01-23T12:34:56-05:00</td>
</tr>
<tr>
<td>degradedCelldisks</td>
<td></td>
</tr>
<tr>
<td>effectiveSize</td>
<td>512M</td>
</tr>
<tr>
<td>efficiency</td>
<td>100.0</td>
</tr>
<tr>
<td>id</td>
<td>8a0aadc84-908804c4e08e1c-b6bcbd5c1ba</td>
</tr>
<tr>
<td>size</td>
<td>512M</td>
</tr>
<tr>
<td>status</td>
<td>normal</td>
</tr>
</tbody>
</table>

8.7.12.12 LIST GRIDDISK

Purpose

The LIST GRIDDISK command displays attributes for one or more Oracle Exadata Storage Server grid disks determined by the specified attributes and filters.

Syntax

LIST GRIDDISK [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-99.

The asmDeactivationOutcome attribute can be used to determine if a grid disk can be deactivated without loss of data. This attribute is not included in the list of attributes shown by the DESCRIBE GRIDDISK command. When using this attribute, a YES in the output means the grid disk can be deactivated.

The asmModeStatus attribute can be used to determine the current usage of a grid disk. This attribute is not included in the list of attributes shown by the DESCRIBE GRIDDISK command. The possible values for this attribute are as follows:
ONLINE: Oracle ASM is actively using this grid disk.
OFFLINE: Oracle ASM has taken this grid disk offline.
DROPPED: Oracle ASM has dropped this grid disk.
UNUSED: No Oracle ASM instance has used this grid disk on the storage cell.
SYNCING: Oracle ASM has started to set this grid disk to online.
UNKNOWN: Oracle ASM instances that use the grid disk are not available to query, or Oracle ASM has rejected the query because it is not in a currently-mounted disk group.

When the cachingPolicy attribute is set to none, the associated flash cache is used for write I/O latency capping and logging. It is not used for caching.

Examples

Example 8-165
Example 8-166
Example 8-167
Example 8-168
Example 8-169
Example 8-170

Example 8-165    Listing Grid Disk Attributes

This example shows the LIST command with the GRIDDISK object, and the corresponding output.

CellCLI> LIST GRIDDISK WHERE cellDisk = 'CD_01_cell05' -
            ATTRIBUTES name, status

            DATA_CD_01_cell105        active
            RECO_CD_01_cell105        active

CellCLI> LIST GRIDDISK DATA_CD_01_cell105 DETAIL

            name:                   DATA_CD_01_cell105
            status:                 active
            comment:                
            id:                     00000117-84d9-0096-0000-000000000000
            creationTime:           2009-01-16T17:04:49-06:00
            cellDisk:               CD_01_cell105
            offset:                 0
            availableTo:            CLUSTER-C1
            size:                   10G
            errorCount:             0
            diskType:               HardDisk
            cachedBy:               FD_01_FLASH, FD02_FLASH, FD03_FLASH
            cachingPolicy:          default

CellCLI> LIST GRIDDISK DATA_CD_01_cell105 ATTRIBUTES size

            136.640625G

CellCLI> LIST GRIDDISK WHERE status!=active ATTRIBUTES name
Example 8-166  Determining if a Grid Disk can be Deactivated

This example shows the `asmDeactivationOutcome` attribute being used to determine if a grid disk can be deactivated.

```
CellCLI> LIST GRIDDISK ATTRIBUTES name, asmDeactivationOutcome

QUAL_CD_00_cell01       Yes
PROD_CD_02_cell01       Cannot de-activate due to other offline disks in the diskgroup
TEST_CD_03_cell01       Yes
DATA_CD_04_cell01       Yes
DATA_CD_05_cell01       Yes
DATA_CD_06_cell01       Yes
RECO_CD_01_cell01       Cannot de-activate due to other offline disks in the diskgroup
DATA_CD_08_cell01       Yes
DATA_CD_09_cell01       Yes
DATA_CD_10_cell01       Yes
DATA_CD_11_cell01       Yes
```

Example 8-167  Viewing the Current Usage of a Grid Disk

This example shows the `asmModeStatus` attribute being used to check the current usage of a grid disk.

```
CellCLI> LIST GRIDDISK ATTRIBUTES name, asmModeStatus

QUAL_CD_00_cell01        UNUSED
RECO_CD_01_cell01        OFFLINE
PROD_CD_02_cell01        SYNCING
TEST_CD_03_cell01        UNKNOWN
DATA_CD_04_cell01        ONLINE
DATA_CD_05_cell01        ONLINE
DATA_CD_06_cell01        ONLINE
```
Example 8-168  Checking the Status Using the LIST GRIDDISK Command

This example shows the LIST GRIDDISK command being used to check the status of the resize process when the NOWAIT option is used with the ALTER GRIDDISK command.

CellCLI> LIST GRIDDISK DETAIL

name:                   gd0
availableTo:           c9standby0
comment:               
creationTime:          2009-07-09T09:07:36-07:00
diskType:              HardDisk
errorCount:            0
id:                    00000122-6045-173b-0000-000000000000
resizeStatus:          Resize in progress
offset:                48M
size:                  48M
status:                active

Example 8-169  Checking the Status of Secure Erase

This example shows the LIST GRIDDISK command being used to check the status of grid disks that are being erased.

CellCLI> LIST GRIDDISK
DATA_CD_00_cell01     active
DATA_CD_05_cell01     active
DATA_CD_06_cell01     erase in progress
DATA_CD_07_cell01     erase in progress

Example 8-170  Listing Devices Used to Cache Grid Disks

This example shows the devices that are currently being used to cache the grid disks.

CellCLI> LIST GRIDDISK ATTRIBUTES name, diskType, size, cachedby

C_DATA_CD_00_dm01celadm01   HardDisk  3.41796875T "FD_02_dm01celadm01, PMEM_02_dm01celadm01"
C_DATA_CD_01_dm01celadm01   HardDisk  3.41796875T "FD_01_dm01celadm01, PMEM_00_dm01celadm01"
C_DATA_CD_02_dm01celadm01   HardDisk  3.41796875T "FD_00_dm01celadm01, PMEM_04_dm01celadm01"
C_DATA_CD_03_dm01celadm01   HardDisk  3.41796875T "FD_02_dm01celadm01, PMEM_01_dm01celadm01"
Purpose

The LIST IBPORT command displays attributes for InfiniBand ports determined by the specified attributes and filters.

Syntax

LIST IBPORT  [ name | attribute_filters ]  [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-100.

If the activeSlave attributes for both InfiniBand ports on a server are not listed, then active-active bonding is being used. If an activeSlave attribute is TRUE, then active-passive bonding is being used.

Examples

The following example shows the LIST command with the IBPORT object, and the corresponding output.

Example 8-171  Listing IBPORT Attributes

CellCLI> LIST IBPORT
    HCA-1:1       Active
    HCA-1:2       Active

CellCLI> LIST IBPORT DETAIL
    name:            HCA-1:1
    activeSlave:     TRUE
    dataRate:        "40 Gbps"
    hcaFWVersion:    2.7.0
    id:              0x00212800013e8c67
    lid:             20
    linkDowned:      0
    linkIntegrityErrs: 0
    linkRecovers:    0
    physLinkState:   LinkUp
    portNumber:      1
    rcvConstraintErrs: 0
    rcvData:         84653709
    rcvErrs:         0
    rcvRemotePhysErrs: 0
    status:          Active
    symbolErrs:      0
    vl15Dropped:     0
    xmtConstraintErrs: 0
    xmtData:         84572496
    xmtDiscards:     0

    name:            HCA-1:2
    activeSlave:     FALSE
    dataRate:        "40 Gbps"
The LIST IORMPLAN command lists the current plan of the local cell.

Syntax

LIST IORMPLAN [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-101.

Examples

The following example shows the LIST command with the IORMPLAN object, and the corresponding output.

Example 8-172  Listing IORMPLAN Attributes

CellCLI> LIST IORMPLAN ATTRIBUTES status

active

CellCLI> LIST IORMPLAN DETAIL

name: cell01_IORMPLAN
catPlan: name=administrative,level=1,allocation=80
         name=interactive,level=2,allocation=90
         name=batch,level=3,allocation=80
         name=maintenance,level=4,allocation=50
         name=other,level=4,allocation=50
dbPlan: name=sales_prod,level=1,allocation=80
         name=finance_prod,level=1,allocation=20
         name=sales_dev,level=2,allocation=100
         name=sales_test,level=3,allocation=50
         name=other,level=3,allocation=50
8.7.12.15 LIST IORMPROFILE

**Purpose**

The **LIST IORMPROFILE** command enables you to list IORM profiles. To see which databases are associated with an IORM profile, use the **LIST DATABASE** command with the new **PROFILE** attribute.

**Syntax**

```
LIST IORMPROFILE [name | filters]
```

**Usage Notes**

- `name` specifies the IORM profile to display.
- `filters` specifies an expression that determines which IORM profiles to display.

If `name` and `filters` are omitted, all the IORM profiles are displayed.

**Examples**

The following example shows the **LIST IORMPROFILE** command, and the corresponding output.

**Example 8-173**  **LIST IORMPROFILE**

```
CellCLI> LIST IORMPROFILE
GOLD
SILVER
BRONZE
```

**Related Topics**

- [About I/O Resource Management Profiles](#)
  I/O Resource Management (IORM) interdatabase plans support profiles to ease management, and configuration of interdatabase plans for hundreds of databases.

8.7.12.16 LIST KEY

**Purpose**

The **LIST KEY** command displays key values for clients determined by the specified attributes and filters.

**Syntax**

```
LIST KEY [name | filters] [attribute_list] [DETAIL]
```

```bash
LIST KEY [FOR {ASMCLUSTER | CELL | LOCAL CELL | REMOTE CELL}] [DETAIL]
```
Usage Notes

• The key value assigned to a client must match the keys in the cellkey.ora files on cells, and the Oracle ASM and database host computers.

• The output of the command might show a type attribute, which was introduced in Oracle Exadata System Software release 12.2.1.1.0. Values for this attribute include ASMCLUSTER, LOCAL-CELL, REMOTE-CELL, and CELL. See the examples below.

• The FOR [LOCAL | REMOTE] CELL and FOR ASMCLUSTER clauses were also introduced in Oracle Exadata System Software release 12.2.1.1.0. These clauses indicate that the command is to display only the keys with the specified type.

• name specifies the key to display.

Examples

The following example shows the LIST command with the KEY object, and the corresponding output.

Example 8-174  Listing KEY Attributes

```
CellCLI> LIST KEY db1 DETAIL
name: db1
key: b67d5587fe728118af47c57ab8da650a

CellCLI> LIST KEY
db1  b67d5587fe728118af47c57ab8da650a
db456 118af47c57ab8da650ab67d5587fe728
asm1 118af47c57ab8da650ab67d5587fe728  ASMCLUSTER

CellCLI> LIST KEY asml DETAIL
name: asml
key: b67d5587fe728118af47c57ab8da650a
type: ASMCLUSTER

CellCLI> LIST KEY FOR CELL DETAIL
key: fa292e11b31b210c4b7a24c5f1bb4d32
  type: CELL
```

Related Topics

• About Security Keys

8.7.12.17 LIST LUN

Purpose

The LIST LUN command displays attributes for LUNs determined by the specified attributes and filters.
Syntax

LIST LUN [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-103.

Examples

The following example shows the LIST command with the LUN object, and the corresponding output.

Example 8-175  Listing LUN Attributes

CellCLI> LIST LUN
  0_0     0_0     normal
  0_1     0_1     normal
  0_2     0_2     normal
  0_3     0_3     normal
  0_4     0_4     normal
  0_5     0_5     normal
  0_6     0_6     normal
  0_7     0_7     normal
  0_8     0_8     normal
  0_9     0_9     normal
  0_10    0_10    normal
  0_11    0_11    normal
  1_0     1_0     normal
  1_1     1_1     normal
  1_2     1_2     normal
  1_3     1_3     normal
  2_0     2_0     normal
  2_1     2_1     normal
  2_2     2_2     normal
  2_3     2_3     normal
  4_0     4_0     normal
  4_1     4_1     normal
  4_2     4_2     normal
  4_3     4_3     normal
  5_0     5_0     normal
  5_1     5_1     normal
  5_2     5_2     normal
  5_3     5_3     normal

CellCLI> LIST LUN 0_0 DETAIL
  name:                   0_0
  cellDisk:               CD_00_sgsas1
  deviceName:             /dev/sda
  diskType:               HardDisk
  id:                     0_0
  isSystemLun:            TRUE
  lunAutoCreate:          TRUE
  lunSize:                558.40625G
8.7.12.18 LIST METRICCURRENT

Purpose

The LIST METRICCURRENT command displays a list of collections of all metrics.

Syntax

LIST METRICCURRENT [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

• The list of attributes that can be displayed is shown in Example 8-104.
• To reduce the size of the output when you run the LIST METRICCURRENT command, use filters.

Examples

Example 8-176 shows the LIST command with filters to display information about the METRICCURRENT object, and the corresponding output.

Example 8-177 shows the LIST METRICCURRENT command with the ORDER BY and LIMIT options.

Example 8-176  Listing METRICCURRENT Attributes

CellCLI> LIST METRICCURRENT WHERE objectType = 'CELLDISK'

| Name               | Object      | Time  
|--------------------|-------------|-------
| CD_IO_TM_W_SM_RQ   | c9controlfile0 | 205.5 us/request 
| CD_IO_TM_W_SM_RQ   | c9datafile0   | 93.3 us/request   
| CD_IO_TM_W_SM_RQ   | c9datafile1   | 0.0 us/request    
| CD_IO_TM_W_SM_RQ   | c9datafile2   | 110.5 us/request  
| CD_IO_TM_W_SM_RQ   | c9datafile3   | 0.0 us/request    
| CD_IO_TM_W_SM_RQ   | c9datafile4   | 541.5 us/request  
| CD_IO_TM_W_SM_RQ   | c9logfile0    | 181.2 us/request  

Chapter 8

CellCLI Command Reference
CD_IO_TM_W_SM_RQ  c9logfile1  0.0 us/request
CD_IO_TM_W_SM_RQ  c9standby0  130.4 us/request

CellCLI> LIST METRICCURRENT WHERE name = CD_IO_TM_W_SM_RQ   -
AND metricObjectName = c9datafile4 DETAIL

name:               CD_IO_TM_W_SM_RQ
alertState:         normal
collectionTime:     2009-07-01T15:19:25-07:00
metricObjectName:   c9datafile4
metricType:         Rate
metricValue:        0.0 us/request
objectType:         CELLDISK

CellCLI> LIST METRICCURRENT CG_IO_UTIL_LG

CG_IO_UTIL_LG
RDB1.BATCH_GROUP            0 %
CG_IO_UTIL_LG
RDB1.INTERACTIVE_GROUP      0 %
CG_IO_UTIL_LG
RDB1.OTHER_GROUPS           0 %
CG_IO_UTIL_LG RDB2.BATCH_GROUP 0 %
CG_IO_UTIL_LG RDB2.INTERACTIVE_GROUP 0 %
CG_IO_UTIL_LG RDB2.OTHER_GROUPS 0 %

Example 8-177  Listing METRICCURRENT Attributes Using ORDER BY and LIMIT

CellCLI> LIST METRICCURRENT attributes name, metricObjectName,
alertState,  \
metricValue ORDER BY metricValue desc, metricObjectName
asc,  \
name desc LIMIT 3

CD_IO_TM_R_LG  c9FLASH0 normal  160,514,088 us
CD_IO_TM_R_LG  c9FLASH1 normal  156,659,463 us
DB_IO_TM_SM   ASM        normal  33,111,890 us

Related Topics

- **Attribute Filters in LIST and ALTER GRIDDISK Commands**
  You can specify which objects to display for the LIST and ALTER GRIDDISK commands with the optional attribute_filters clause.

- **Exadata Metrics**
  Exadata metrics are recorded observations of important properties or values relating to the Exadata system software.
8.7.12.19 LIST METRICDEFINITION

Purpose

The LIST METRICDEFINITION command displays a list of metric definitions on the cell.

Syntax

LIST METRICDEFINITION [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-105.

Examples

The following example shows the LIST command with the METRICDEFINITION object, and the corresponding output.

Example 8-178  Listing the Metric Definitions for a Specific Object

CellCLI> LIST metricDefinition WHERE objectType=cell
CL_BBU_CHARGE
CL_BBU_TEMP
CL_CPUT
CL_CPUT_CS
CL_CPUT_MS
CL_FANS
CL_IO_RQ_NODATA
CL_IO_RQ_NODATA_SEC
CL_MEMUT
CL_MEMUT_CS
CL_MEMUT_MS
CL_RUNQ
CL_SWAP_IN_BY_SEC
CL_SWAP_OUT_BY_SEC
CL_SWAP_USAGE
CL_TEMP
CL_VIRTMEM_CS
CL_VIRTMEM_MS
IORM_MODE
N_HCA_MB_RCV_SEC
N_HCA_MB_TRANS_SEC
N_NIC Kb_RCV_SEC
N_NIC Kb_TRANS_SEC

Example 8-179  Listing the Metric Definition Detail for a Specific Metric

CellCLI> LIST metricDefinition WHERE name=cl_swap_in_by_sec DETAIL
name: CL_SWAP_IN_BY_SEC
description: "Amount of swap pages read in KB per second"
Related Topics

- **Exadata Metrics**
  Exadata metrics are recorded observations of important properties or values relating to the Exadata system software.

8.7.12.20 LIST METRICHISTORY

**Purpose**

The `LIST METRICHISTORY` command displays a list of individual metrics.

**Syntax**

```
LIST METRICHISTORY [ name | attribute_filters ] [attribute_list]
   [over_specification] [MEMORY] [DETAIL]
```

**Usage Notes**

- The list of attributes that can be displayed is shown in Example 8-106.
- The retention period for metric history entry files is specified by the `metricHistoryDays` cell attribute. You can modify this setting with the `CellCLI ALTER CELL` command.
- The `over_specification` syntax is as follows:

  ```
  OVER number [aggregation_type [aggregation_type]...]
  ```

  In the preceding syntax, `number` is amount of time in minutes for the aggregation, and `aggregation_type` can be `max`, `min`, or `avg`.
- A WHERE clause can include the `ageInMinutes` attribute to specify the list is limited to those metrics which have the specified age. For example, the following command would show the metrics created in the previous 15 minutes:

  ```
  CellCLI> LIST METRICHISTORY WHERE ageInMinutes < 15
  ```

**Examples**

**Example 8-180  Listing METRICHISTORY Attributes**

This example shows the `LIST` command with the `METRICHISTORY` object, and the corresponding output. To reduce the size of the output when you run the `LIST METRICHISTORY` command, use filters.

```
CellCLI> LIST METRICHISTORY WHERE name like 'CL_.*' -
     AND collectionTime > '2009-07-01T15:28:36-07:00'

   CL_RUNQ         stbcr03_2       6.0
   2009-07-01T15:28:37-07:00
```
Example 8-181  Listing METRICHISTORY Using the OVER and MEMORY Attributes

This example shows the LIST METRICHISTORY command with the OVER and MEMORY options.

```
Example 8-181    Listing METRICHISTORY Using the OVER and MEMORY Attributes

This example shows the LIST METRICHISTORY command with the OVER and MEMORY options.

CellCLI> LIST METRICHISTRY cl_cput OVER 10 MIN MAX MEMORY

CL_CPUT         firstcell       55.0 %  2009-11-15T06:00:17-08:00      55.0
% 57.1 %
CL_CPUT         firstcell       54.7 %  2009-11-15T06:10:17-08:00      54.7
% 56.3 %
CL_CPUT         firstcell       54.8 %  2009-11-15T06:20:18-08:00      54.7
% 57.2 %
CL_CPUT         firstcell       55.0 %  2009-11-15T06:30:18-08:00      54.3
% 55.9 %
CL_CPUT         firstcell       55.0 %  2009-11-15T06:40:18-08:00      54.9
% 57.0 %
CL_CPUT         firstcell       55.1 %  2009-11-15T06:50:18-08:00      54.8
% 56.4 %
CL_CPUT         firstcell       58.0 %  2009-11-15T07:00:18-08:00      55.2
% 58.0 %
CL_CPUT         firstcell       55.5 %  2009-11-15T07:10:18-08:00      55.5
% 67.5 %
```

Example 8-182  Listing METRICHISTORY Attributes Using ORDER BY and LIMIT

This example shows the LIST METRICHISTRY command with the ORDER BY and LIMIT options.

```
Example 8-182    Listing METRICHISTORY Attributes Using ORDER BY and LIMIT

This example shows the LIST METRICHISTRY command with the ORDER BY and LIMIT options.

CellCLI> LIST METRICHISTRY WHERE name like '.IO_RQ.*' DETAIL ORDER

name:                   CT_FD_IO_RQ_SM
alertState:             normal
metricObjectName:       OTHER
metricType:             Cumulative
metricValue:            3,211,568 IO requests
objectType:             IORM_CATEGORY

name:                   CT_FD_IO_RQ_SM
alertState:             normal
collectionTime:         2014-05-23T10:58:06-07:00
metricObjectName:       OTHER
metricType:             Cumulative
```
Example 8-183  Listing METRICHISTORY Attributes with ASM-scoped security configured

This example for Oracle Exadata System Software release 19.1.0 shows the LIST METRICHISTORY command with the DETAIL option for a system with ASM-scoped security configured.

CellCLI> LIST METRICHISTORY WHERE name like 'DB_IO_RQ_SM' DETAIL ORDER BY -
metricObjectName LIMIT 4

name: DB_IO_RQ_SM
alertState: normal
collectionTime: 2018-08-23T07:34:12-05:00
metricObjectName: ASM
metricType: Cumulative
metricValue: *
objectType: IORM_DATABASE

name: DB_IO_RQ_SM
alertState: normal
collectionTime: 2018-08-23T07:33:38-05:00
metricObjectName: ASM1.PRODDB
metricType: Cumulative
metricValue: *
objectType: IORM_DATABASE

name: DB_IO_RQ_SM
alertState: normal
collectionTime: 2018-08-22T14:27:45-05:00
metricObjectName: ASM1.PRODDB
metricType: Cumulative
metricValue: *
objectType: IORM_DATABASE

name: DB_IO_RQ_SM
**8.7.12.21 LIST OFFLOADGROUP**

**Purpose**

The LIST OFFLOADGROUP command displays the attributes for offload groups.

**Syntax**

```
LIST OFFLOADGROUP [name | filters] [attribute_list] [DETAIL]
```

**Usage Notes**

- The *name* and *filters* parameters specify the offload groups for which you want to display the attributes.
  - *name* specifies the name of an offload group.
  - *filters* specifies an expression to match one or more offload groups.
  - If neither *name* nor *filters* is specified, the command lists the attributes for all offload groups.

- The *attribute_list* parameter specifies one or more attributes for which you want to view. The *attribute_list* begins with the ATTRIBUTES keyword. If specifying more than one attribute, separate the attributes with a comma:
  
  `ATTRIBUTES { attr1 [, attr2]... }`

  If the *attribute_list* parameter is omitted, the command displays all attributes.

- If the DETAIL parameter is used, the command displays an attribute descriptor for each attribute.

**Examples**

**Example 8-184  Displaying All Offload Groups in Detail**

The following command lists all the attributes for all offload groups.

```
LIST OFFLOADGROUP DETAIL
```
Example 8-185  Displaying the Attributes for an Offload Group

The following command lists all the attributes for an offload group named "offloadgroup1".

LIST OFFLOADGROUP offloadgroup1

Example 8-186  Displaying Specific Attributes

The following command lists the name and package attributes for all offload groups.

LIST OFFLOADGROUP ATTRIBUTES name, package

Related Topics

• ALTER OFFLOADGROUP

8.7.12.22 LIST OFFLOADPACKAGE

Purpose

The LIST OFFLOADPACKAGE command displays the attributes for offload packages. The offload packages provide information on the offload server versions that are installed on the cell. This indirectly shows the various database versions that are supported on the cell.

Syntax

LIST OFFLOADPACKAGE [name | filters] [attribute_list] [DETAIL]

Usage Notes

• The name and filters parameters specify the offload packages for which you want to display the attributes.
  – name The name of the offload package to be displayed.
  – filters an expression which determines which offload packages should be displayed.
  – If neither name nor filters is specified, the command lists the attributes for all offload packages.

• The attribute_list parameter specifies one or more attributes for which you want to view. The attribute_list begins with the ATTRIBUTES keyword. If specifying more than one attribute, separate the attributes with a comma:

  ATTRIBUTES ( attr1 [, attr2]... )

  If the attribute_list parameter is omitted, the command displays all attributes.

• If the DETAIL parameter is used, the command displays an attribute descriptor for each attribute.

Examples

Example 8-187  Displaying All Offload Packages in Detail

The following command lists all the attributes for all offload packages.

LIST OFFLOADPACKAGE DETAIL
Example 8-188  Displaying the Attributes for an Offload Package

The following command lists all the attributes for an offload package named 'cellofl-12.1.1.1.1_LINUX.X64_130211'.

LIST OFFLOADPACKAGE 'cellofl-12.1.1.1.1_LINUX.X64_130211'

Example 8-189  Displaying Specific Attributes

The following command lists the name and ispublic attributes for all offload packages.

LIST OFFLOADPACKAGE ATTRIBUTES name, ispublic

Example 8-190  Displaying Packages with Specific Attribute Values

The following command lists the name of all packages where the ispublic attribute.

LIST OFFLOADPACKAGE where ispublic = 'true'

CellCLI> describe offloadpackage
    name
    installationTime
    isPublic

CellCLI> list offloadpackage
    cellofl-12.1.2.4.0_LINUX.X64_200526
    cellofl-21.1.0.0.0_LINUX.X64_200912
    cellofl-11.2.3.3.1_LINUX.X64_200526

Related Topics
• ALTER OFFLOADGROUP

8.7.12.23 LIST PHYSICALDISK

Purpose

The LIST PHYSICALDISK command displays attributes for one or more physical disks determined by the specified attributes and filters.

Syntax

LIST PHYSICALDISK [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-108.

When a physical disk is performing a power cycle, the status of the disk is included in the output. The status options are Normal-DiskPoweredOn, Normal-DiskPoweredOff, ProactiveFailure-DiskPoweredOn, and ProactiveFailure-
DiskPoweredDown. If physical disk status shows as failed, perform a power cycle on the disk to verify the status of the physical disk.

Example 8-191 Listing Physical Disk Attributes

This example shows the LIST command with the PHYSICALDISK object, and the corresponding output.

```
CellCLI> LIST PHYSICALDISK
20:0            K68DWJ          normal
20:1            K7YXUJ          normal
20:2            K7TYEJ          normal
20:3            K7BJMJ          normal
20:4            K5B4SM          normal
20:5            KEBTDJ          normal
20:6            K4URJJ          normal
20:7            K5E1DM          normal
20:8            K7VL6J          normal
20:9            K7N5NJ          normal
20:10           K723KJ          normal
20:11           K504ZM          normal
FLASH_1_0       1030M03RK1      normal
FLASH_1_1       1030M03RJN      normal
FLASH_1_2       1030M03RJH      normal
FLASH_1_3       1030M03RJD      normal
FLASH_2_0       1027M03N6X      normal
FLASH_2_1       1027M03N6Y      normal
FLASH_2_2       1027M03N6W      normal
FLASH_2_3       1027M03N6Y      normal
FLASH_4_0       1025M03EJ3      normal
FLASH_4_1       1025M03EJ2      normal
FLASH_4_2       1025M03EJ2      normal
FLASH_4_3       1025M03EJ2      normal
FLASH_5_0       1028M03QP8      normal
FLASH_5_1       1028M03QN0      normal
FLASH_5_2       1028M03QKU      normal
FLASH_5_3       1028M03QH4      normal

CellCLI> LIST PHYSICALDISK 20:0 DETAIL
name:                   20:0
deviceId:               46
deviceName:             /dev/sda
diskType:               HardDisk
enclosureDeviceId:      20
errOtherCount:          0
luns:                   0_0
makeModel:              "HITACHI HUS1560SCSUN600G"
physicalFirmware:       A8C0
physicalInsertTime:     2017-07-27T07:03:00-04:00
physicalInterface:      sas
physicalSerial:         K68DWJ
physicalSize:           558.9120712280273G
slotNumber:             0
status:                 normal
```
CellCLI> LIST PHYSICALDISK FLASH_5_3 DETAIL

name:                   FLASH_5_3
deviceName:             /dev/sdx
diskType:               FlashDisk
luns:                   5_3
makeModel:              "Sun Flash Accelerator F20 PCIe Card"
physicalFirmware:       D21Y
physicalInsertTime:     2017-07-27T07:03:01-04:00
physicalSerial:         1028M03QHN
physicalSize:           22.8880615234375G
slotNumber:             "PCI Slot: 5; FDOM: 3"
status:                 normal

Related Topics

• ALTER PHYSICALDISK
• About Quoting Object Names
  If an object name has characters that cause parsing errors enclose the object name in quotes.

8.7.12.24 LIST PLUGGABLEDATABASE

Purpose
Displays the specified attributes for active pluggable databases.

Syntax

LIST PLUGGABLEDATABASE [name | attribute_filters] [attribute_list]
[DETAIL]

Usage Notes
The filters option is an expression that determines which active pluggable databases should be listed by the command.

Examples
The following example shows the LIST PLUGGABLEDATABASE command and the corresponding output.

Example 8-192    Listing Pluggable Database Attributes

CellCLI> LIST PLUGGABLEDATABASE
  PDB$SEED
  CDB1_PDB1
  NEWPDB1
  NEWPDB2

CellCLI> LIST PLUGGABLEDATABASE DETAIL
  name:                   PDB$SEED
  asmClusterName:         SALESDBS_ASMCLUSTER
  containerName:          CDB1
  flashCacheLimit:        515M
<table>
<thead>
<tr>
<th>Name</th>
<th>asmClusterName</th>
<th>ContainerName</th>
<th>Sample Limit</th>
<th>Sample Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDB1_PDB1</td>
<td>SALESDBS_ASMCLUSTER</td>
<td>CDB1</td>
<td>315G</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200G</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>NEWPDB1</td>
<td>SALESDBS_ASMCLUSTER</td>
<td>CDB1</td>
<td>157G</td>
<td>80.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20G</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>NEWPDB2</td>
<td>SALESDBS_ASMCLUSTER</td>
<td>CDB1</td>
<td>157G</td>
<td>60.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20G</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
8.7.12.25 LIST PMEMCACHE

Purpose
The LIST PMEMCACHE command displays the specified attributes for the PMEM cache.

Syntax
LIST PMEMCACHE [attribute_list] [DETAIL]

Usage Notes
The list of attributes that can be displayed is shown in DESCRIBE PMEMCACHE.

Examples
The following example shows the LIST command with the PMEMCACHE object, and the corresponding output.

Example 8-193  Listing PMEM Cache Attributes

CellCLI> LIST PMEMCACHE
dbm01celadm08_PMEMCACHE  normal

CellCLI> LIST PMEMCACHE DETAIL
  name:       dbm01celadm08_PMEMCACHE  
  cellDisk:   PM_10_dbm01celadm08,PM_11_dbm01celadm08,PM_06_dbm01celadm08,PM_01_dbm01celadm08,PM_00_dbm01celadm08,PM_03_dbm01celadm08,PM_08_dbm01celadm08,PM_02_dbm01celadm08,PM_09_dbm01celadm08,PM_07_dbm01celadm08,PM_04_dbm01celadm08,PM_05_dbm01celadm08
  creationTime:  2019-09-12T11:37:00-07:00
  degradedCelldisks: 
  effectiveCacheSize:  1.474365234375T
  id:       0ba69a26-b02d-46a8-a5a6-82e699a6ac88
  size:     1.474365234375T
  status:   normal

8.7.12.26 LIST PMEMLOG

The LIST PMEMLOG command displays the specified attributes for the PMEMLOG.

Syntax
LIST PMEMLOG [attribute_list] [DETAIL]

Usage Notes
The list of attributes that can be displayed is shown in DESCRIBE PMEMLOG.
Examples

The following example shows the LIST command with the PMEMLOG object, and the corresponding output.

**Example 8-194  Listing the Current PMEMLOG**

The following example shows the LIST PMEMLOG command with the default values.

```
CellCLI> LIST PMEMLOG

raw_PMEMLOG    normal
```

**Example 8-195  Listing All PMEMLOG Attributes**

```
CellCLI> LIST PMEMLOG DETAIL

name:                   raw_PMEMLOG
cellDisk:               NV_00_raw,NV_01_raw,NV_02_raw,NV_03_raw
creationTime:           2019-11-23T12:34:56-05:00
degradedCelldisks:     
effectiveSize           1G
efficiency:             100.0
id:                     8a0adc84-9088-4c4e-8e1c-b6bcbd5cb1ba
size:                   1G
status:                 normal
```

8.7.12.27 LIST QUARANTINE

**Purpose**

The LIST QUARANTINE command displays specified attributes for quarantines.

**Syntax**

```
LIST QUARANTINE [ name | attribute_filters ] [attribute_list] [DETAIL]
```

**Examples**

The following example shows the LIST command with the QUARANTINE object.

**Example 8-196  Listing QUARANTINE Attributes**

```
CellCLI> LIST QUARANTINE DETAIL

CellCLI> LIST QUARANTINE where comment like 'added.*'
```

**Related Topics**

- DESCRIBE QUARANTINE
8.7.12.28 LIST ROLE

Purpose

The LIST ROLE command displays the specified attributes for a role.

Syntax

LIST ROLE [name | filters] [ATTRIBUTES {ALL | attr1 [,attr2]...}] [DETAIL]

Usage Notes

• Use name to display information about a specific role.
• Use an expression in place of filters to display information about the roles that satisfy the expression.
• Use the ATTRIBUTES keyword to display information about one or more attributes of the roles. The ALL option can be used to display all attributes.
• Use the keyword DETAIL to format the output as an attribute on each line, with an attribute descriptor preceding each value.

Example 8-197  Displaying Roles

This example shows how to display detailed information about all roles.

CellCLI> LIST ROLE DETAIL
    name: admin
    privileges: object=all objects, verb=all actions,
    attributes=all attributes, options=all options

    name: gd_monitor
    privileges: object=griddisk, verb=0,
    attributes=all attributes, options=all options

Example 8-198  Displaying Roles Using an Expression

This example shows how to display roles using an expression.

CellCLI> LIST ROLE WHERE name>'ad' AND name<'ba'
    admin

8.7.12.29 LIST SOFTWAREHISTORY

Purpose

The LIST SOFTWAREHISTORY command displays a list of final states for past software updates.
Syntax

LIST SOFTWAREHISTORY [attribute_filters] [attribute_list] [DETAIL]

Example 8-199  Displaying the History of a Scheduled Software Update

By default, only the update name and status are shown.

CellCLI> LIST SOFTWAREHISTORY
12.2.1.2.0.170509  Last update completed at: 2017-05-20T08:00:57-07:00
12.2.1.2.0.170520  Last update completed at: 2017-05-21T06:39:54-07:00
12.2.1.2.0.17052   Last update completed at: 2017-06-08T08:56:45-07:00
12.2.1.2.0.170603  Last update completed at: 2017-06-08T16:03:17-07:00

Example 8-200  Displaying the Detailed History of a Specific Software Update

This example shows the detailed software update history for a particular update which is referenced by its name.

CellCLI> LIST SOFTWAREHISTORY WHERE name='12.2.1.2.0.170808.1' DETAIL
name:                   12.2.1.2.0.170808.1
status:                 Upgrade failed. See alerts at:
2017-08-10T10:56:15-07:00

8.7.12.30 LIST SOFTWAREUPDATE

Purpose

The LIST SOFTWAREUPDATE command displays the status of the most recently scheduled update.

Syntax

LIST SOFTWAREUPDATE [attribute_list] [DETAIL]

Usage Notes

The possible states for the software update are:

• Ready to update at: update_time
• Downloading
• Checking prerequisites
• Prerequisites failed. See alerts.
• Last update completed at: update_time
• Running
• Upgrade failed. See alerts.
Example 8-201  Displaying the Status of a Scheduled Software Update

By default, only the update name and status are shown.

CellCLI> LIST SOFTWAREUPDATE
    12.2.1.2.0.170603   Last update completed at: 2017-06-08 16:03:17 -0700

Example 8-202  Displaying the Detailed Status of a Scheduled Software Update

By default, only the update name and status are shown.

CellCLI> LIST SOFTWAREUPDATE DETAIL
    name:           12.2.1.2.0.170808.1
    status:         Upgrade failed. See alerts
    store:          https://mystore_url:4443
    time:           2017-08-10T10:35:00-07:00

8.7.12.31 LIST THRESHOLD

Purpose

The LIST THRESHOLD command displays attributes for one or more thresholds determined by the specified attributes and filters.

Syntax

LIST THRESHOLD [ name | attribute_filters ] [attribute_list] [DETAIL]

Usage Notes

The list of attributes that can be displayed is shown in Example 8-116.

Examples

The following example shows the LIST command with the THRESHOLD object, and the corresponding output.

Example 8-203  Listing Threshold Attributes

CellCLI> LIST THRESHOLD
    ct_io_wt_rq.interactive
db_io_rq_sm_sec.db123
    ....

CellCLI> LIST THRESHOLD ct_io_wt_rq.interactive DETAIL

    comparison:     =
    critical:       20.0
    name:           ct_io_wt_rq.interactive
    observation:    5
    occurrences:    2
    warning:        10.0
8.7.12.32 LIST USER

Purpose
The LIST USER command displays the specified attributes for a user.

Syntax
LIST USER [name | filters] [attribute_list] [DETAIL]

Usage Notes
• name is the user name.
• filters is an expression that determines which users are displayed.
• attribute_list is the attributes to display. The ALL option can be used to display all attributes.
• The DETAIL option formats the output as an attribute on each line, with an attribute descriptor preceding each value.

Examples
The following example shows the LIST USER command.

Example 8-204 Using the LIST USER Command

CellCLI> LIST USER DETAIL
CellCLI> LIST USER where name like 'agarcia' DETAIL
name: agarcia
roles: role=gd_monitor
Privileges: object=griddisk verb=list attributes=all attributes options= all options

8.7.13 QUIT

Purpose
The QUIT command exits from the CellCLI utility, and returns control to the operating system prompt.

Syntax
QUIT

QUIT has the same functionality as the EXIT command.
8.7.14 REVOKE

Purpose
The REVOKE command removes privileges and roles.

Syntax
REVOKE object_type [name] FROM sub_object_type [sub_object_name]

Usage Notes
• object_type can be as follows:
  – PRIVILEGE
  – ROLE
• The following can be used for PRIVILEGE object type:
  – name is in the following format:

    ( ALL ACTIONS | action ) ON ( ALL OBJECTS | object ) ( ALL ATTRIBUTES |
    \       ATTRIBUTES attribute1 [, attribute2, ...] ) ( WITH ALL OPTIONS |
    \       WITH OPTIONS option1 [, option2, ...] )

  – The sub_object_type must be ROLE.
  – The sub_object_name is the name of the role.
• The following can be used for the ROLE object type:
  – name is the role name.
  – The sub_object_type must be USER.
  – The sub_object_name is the name of the user.
• REVOKE PRIVILEGE
• REVOKE ROLE

Related Topics
• REVOKE PRIVILEGE
• REVOKE ROLE

8.7.14.1 REVOKE PRIVILEGE

Purpose
The REVOKE PRIVILEGE command revokes privileges from a role.

Syntax
REVOKE PRIVILEGE { ALL ACTIONS | action } ON { ALL OBJECTS | object } { ALL ATTRIBUTES | ATTRIBUTES attribute1 [, attribute2...] } { WITH
ALL OPTIONS | WITH OPTIONS option1 [, option2, ...] } FROM ROLE { ALL | role1 [, role2, ...] } 

Usage Notes
- *action* is the command.
- *object* is object type for the action.
- *attribute* are the attributes for the object.
- *option* are the options for the object.
- *role* is the name of the role from which to revoke privileges.
- The **ALL ACTIONS** argument revokes privileges for all actions.
- The **ALL OBJECTS** argument revokes privileges for all objects.
- The **ALL ATTRIBUTES** argument revokes privileges for all attributes.
- The **WITH ALL OPTIONS** argument revokes privileges for all options.

Examples
The following example shows the **REVOKE PRIVILEGE** command.

**Example 8-205  Revoking a Privilege**

```
CellCLI> REVOKE PRIVILEGE ALL ACTIONS ON ALL OBJECTS ALL ATTRIBUTES -
      WITH ALL OPTIONS FROM ROLE ALL

CellCLI> REVOKE PRIVILEGE list on griddisk ATTRIBUTES name,size
      WITH OPTIONS detail FROM ROLE gd_monitor
```

8.7.14.2 REVOKE ROLE

Purpose
The **REVOKE ROLE** command revokes the role for a user.

Syntax

```
REVOKE ROLE { ALL | role1 [, role2, ...] } FROM USER { ALL | user1 [, user2...] }
```

Usage Notes
- *role* is the name of a role.
- *user* is the name of a user.
- The **ALL** argument revokes all roles from the user.
- The **FROM USER ALL** argument revokes the role from all users.

Examples
The following example shows how to revoke a role from a user.
8.7.15 SET

Purpose
The SET command sets parameter options in the CellCLI environment.

Syntax

```
SET DATEFORMAT {LOCAL | STANDARD}
SET ECHO [ON | OFF]
```

The `SET DATEFORMAT` command controls the format of displayed dates. For commands that accept dates, the standard date-time format is recommended. The local format is also accepted. The standard format is recommended for scripts because that format is less sensitive to the time zone, region, and locale changes that might occur when running a script.

The `SET ECHO` command controls whether to echo commands in a script that is run with `@` or `START`. The `ON` option displays the commands on screen. The `OFF` option suppresses the display. The `SET ECHO` command does not affect the display of commands entered interactively or redirected from the operating system.

Example 8-207  Setting the Date Format with the SET Command

This example shows an example of the `SET` command.

```
SET DATEFORMAT STANDARD
```

8.7.16 SPOOL

Purpose
The SPOOL command writes (spools) the results of commands to the specified file on the cell file system.

Syntax

```
SPOOL [file_name [ CREATE | REPLACE | APPEND ] | OFF]
```

If you issue `SPOOL file_name` with no option, then the output is spooled to that file whether or not the file already exists. The `REPLACE` option is the default behavior.

The SPOOL options are described in the following table.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPEND</td>
<td>Adds the results to the end of the file specified.</td>
</tr>
</tbody>
</table>
### Table 8-4  (Cont.) SPOOL Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CREATE</td>
<td>Creates a new file with the name specified, and raises an error if the file exists.</td>
</tr>
<tr>
<td>file_name</td>
<td>Names the file to which the results are written. It can be specified with a fully-qualified path name, or with a partially-qualified path name relative to the current directory.</td>
</tr>
<tr>
<td>no option</td>
<td>Displays the name of the current spool target file, if any.</td>
</tr>
<tr>
<td>OFF</td>
<td>Stops writing (spooling) output to the file.</td>
</tr>
<tr>
<td>REPLACE</td>
<td>Replaces the contents of an existing specified file. If the file does not exist, then REPLACE creates the file. This is the default behavior.</td>
</tr>
</tbody>
</table>

### 8.7.17 START and @

#### Purpose

The START or @ command runs the CellCLI commands in the specified script file.

#### Syntax

```
START file_name
@file_name
```

The START and @ both require the option `file_name`. It is the name of the script file that contains the CellCLI commands. If the file name does not include a fully-qualified path, then the CellCLI utility searches for the file relative to the current directory.

The START or @ command is useful when entering long or multiple CellCLI commands. For example, the commands in Example 8-3 can be entered in a text file named `alter_cell`, then run with `START alter_cell`, assuming that the `alter_cell` file is in the current directory.
Using the dcli Utility

The dcli utility facilitates centralized management across Oracle Exadata Database Machine.

dcli automates the execution of operating system commands on a set of servers and returns the output to the centralized management location where the dcli utility was run.

- **Overview of the dcli Utility**
  The dcli utility runs operating system commands on multiple servers in parallel threads. The target servers can be Exadata Storage Servers or Exadata database servers. However, the dcli utility does not support an interactive session with a remote application.

- **dcli Syntax**
  This topic describes the syntax for the dcli utility.

- **dcli Examples**
  This section contains examples of the using the dcli utility.

- **Setting Up SSH User-Equivalence on Oracle Exadata Storage Server**
  Setting user-equivalence enables you to issue commands to remote cells without having to enter the password for the cell.

### 9.1 Overview of the dcli Utility

The dcli utility runs operating system commands on multiple servers in parallel threads. The target servers can be Exadata Storage Servers or Exadata database servers. However, the dcli utility does not support an interactive session with a remote application.

You can run the dcli utility on an Exadata Storage Server or Exadata database server, or you can copy the utility to another host computer that acts as a central management server. You can issue a command to be run on multiple servers, or use files that can be copied to servers and then run. The servers are referenced by their host name or IP address.

The dcli utility requires Python version 2.3 or later. You can determine the version of Python by running the `python -V` command. In addition, use of this tool assumes prior setup of SSH user-equivalence to the target servers. You can use the dcli utility initially with the `-k` option to set up SSH user-equivalence. Also, you can manually set up SSH user-equivalence to servers.

Command output (stdout and stderr) is collected and displayed after the copy and command execution is finished on the specified servers. The dcli options allow command output to be abbreviated to minimize non-error output, such as messages showing normal status.
9.2 dcli Syntax

This topic describes the syntax for the dcli utility.

Syntax

dcli [options] [command]

Command Arguments

- **options**: command options
- **command**: Any command that can be run from an operating system prompt.

Command Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>--batchsize=MAXTHDS</td>
<td>Limits the number of parallel execution threads, which limits the number of target servers on which to run the operation in parallel.</td>
</tr>
<tr>
<td>-c cells</td>
<td>Specifies a comma-delimited list of target servers to which commands are sent.</td>
</tr>
<tr>
<td>--ctimeout=CTIMEOUT</td>
<td>Specifies the maximum time in seconds for initial connection to a target server.</td>
</tr>
<tr>
<td>-d destfile</td>
<td>Specifies the target destination directory or file on remote servers to be used when copying files or directories using the -f option.</td>
</tr>
<tr>
<td>-f file</td>
<td>Specifies the files or file template to be copied to the servers. These files are not run. These files can be script files to be run later. The files are copied to the default home directory of the user on the target server.</td>
</tr>
<tr>
<td>-g groupFile</td>
<td>Specifies a file containing a list of target servers to which commands are sent. The servers can be identified by host names or IP addresses.</td>
</tr>
<tr>
<td>-h, --help</td>
<td>Displays help text and then exits.</td>
</tr>
<tr>
<td>--hidestderr</td>
<td>Hide standard error messages (STDERR) for commands run remotely using SSH.</td>
</tr>
<tr>
<td>-k</td>
<td>Sets up SSH user-equivalence for the current user to the servers specified with the -c or -g option by appending public key files to the authorized_keys file on servers.</td>
</tr>
<tr>
<td>-l userId</td>
<td>Identifies the user to log in as on remote servers. The default is the celladmin user.</td>
</tr>
</tbody>
</table>
## Table 9-1  (Cont.) dcli Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-n</td>
<td>Abbreviates nonerror output. Servers that return normal output (return code of 0) only have the server name listed. The -n and -r options cannot be used together.</td>
</tr>
<tr>
<td>-r regexp</td>
<td>Abbreviates the output lines that match a regular expression. All output lines with that pattern are deleted from output, and the servers names from those output lines are listed on one line. The -r and -n options cannot be used together.</td>
</tr>
<tr>
<td>-s sshOptions</td>
<td>Passes a string of options to SSH.</td>
</tr>
<tr>
<td>--scp= scpOptions</td>
<td>Passes a string of options to scp if different from sshoptions.</td>
</tr>
<tr>
<td>--serial</td>
<td>Serializes the process over Oracle Exadata Storage Servers.</td>
</tr>
<tr>
<td>--showbanner,--sh</td>
<td>Show the banner of the remote node when using SSH.</td>
</tr>
<tr>
<td>-t</td>
<td>Displays the target servers that are named with the -c option or in the groupfile identified by the -g option.</td>
</tr>
<tr>
<td>--unkey</td>
<td>Drops keys from the target authorized_keys file on Oracle Exadata Storage Servers.</td>
</tr>
<tr>
<td>-v</td>
<td>Prints the verbose version of messages to stdout.</td>
</tr>
<tr>
<td>--version</td>
<td>Shows the version number of the program and then exits.</td>
</tr>
<tr>
<td>--vmstat=VMSTATOPS</td>
<td>Displays view process, virtual memory, disk, trap, and CPU activity information, depending on the switches.</td>
</tr>
<tr>
<td>-x execFile</td>
<td>Specifies the command file to be copied and run on the servers. The specified file contains a list of commands. A file with the .scl extension is run by the CellCLI utility. A file with a different extension is run by the operating system shell on the server. The file is copied to the default home directory of the user on the target server.</td>
</tr>
</tbody>
</table>

### Usage Notes

For commands that contain punctuation that would be interpreted by the local shell, enclose the command in double quotation marks. If the command includes the following characters, then outer quotation marks and escape characters are required:

- $ (dollar sign)
- ’ (quotation mark)
- < (less than)
- > (greater than)
- ( ) (parentheses)

The backslash (\) is the escape character that allows the characters to be passed to the CellCLI utility without being interpreted by the remote shell.

If the command is complex in terms of punctuation that need escape characters, then it may require that the command be put in a script, and run using the -x option. Within a script, the escape character is not required.
Troubleshooting

If the local dcli process is terminated, then remote commands might continue, but their output and status is unknown.

Return values from the dcli utility are:

- 0: The file or command was copied, and run successfully on all servers.
- 1: One or more servers could not be reached or remote execution returned a nonzero status.
- 2: A local error prevented any command execution.

If any servers are down or do not respond, then a message is written to stderr listing the unresponsive servers. The operations continue on the other servers, and the return code after completion is 1.

9.3 dcli Examples

This section contains examples of the using the dcli utility.

- **Using dcli to Set up SSH User-equivalence for a Current User**
  This example shows how to set SSH user-equivalence for a current user using the -k option.

- **Using dcli with the -n Option**
  This example shows how to run the CellCLI command `ALTER IORMPLAN OBJECTIVE='basic'`, and abbreviates non-error output.

- **Using dcli with the -r Option**
  This example shows how to run the CellCLI command `LIST GRIDDISK`, and deletes the lines in the output that contain `normal`.

- **Using dcli with the -v Option**
  This example shows how to use the verbose (-v) option with SSH.

- **Using dcli with the -t Option**
  This example shows how to use the -t option to list target cells.

- **Using dcli with the -f Option**
  This example shows how to use the -f option.

- **Using dcli with the --vmstat Option**
  This example shows how to use the --vmstat option of dcli.

- **Using dcli with the --hidestderr Option**
  This example shows how to use the --hidestderr option.

- **Using dcli with the --showbanner Option**
  This example shows how to use the --showbanner option.

- **Using dcli to Change an IORM Plan**
  This example shows a CellCLI command that changes the IORMPLAN to active on the target cells in the `mycells` group file.

- **Using dcli with a Script**
  This example shows the CellCLI commands in the `reConfig.scl` file on the target cells.


- **Using dcli to List Grid Disk Status**
  This example shows how to run a CellCLI command that lists the name and status of grid disks on the target cells.

- **Using dcli to List Alert History Information**
  This example shows a CellCLI command that lists alert history name, examined by, severity on the target cells.

- **Using dcli to List Alert History Where examinedby is Not Set**
  This example shows a CellCLI command that lists alert history where examined by has not been set on the target cells.

- **Using dcli to List Current Metric Alert State**
  This example shows a CellCLI command that retrieves metric current objects for the number of MB read in large blocks on a grid disk for a group of cells.

- **Using dcli to List Specific Metric Current Objects in a Group**
  This example shows a CellCLI command that retrieves metric current objects for the number of requests to read or write blocks on a grid disk.

- **Using dcli to List Specific Metric Current Objects**
  This example shows a CellCLI command that lists metric current objects with name equal to `cl_put` (cell CPU utilization) on the target cells.

- **Using dcli to List Physical Disks**
  This example shows a CellCLI command that lists physical disks where status is not equal to normal on the target cells.

- **Using dcli to List Cell Disks with Free Space**
  This example shows a CellCLI command that lists cell disks where free space is less than 100 MB on the target cells.

- **Using dcli to View Alert History**
  This example shows a CellCLI command to view the alert history from a particular period.

9.3.1 Using dcli to Set up SSH User-equivalence for a Current User

This example shows how to set SSH user-equivalence for a current user using the `-k` option.

**Example 9-1  Setting up SSH User-equivalence for a Current User**

```
$ ./dcli -k -g mycells
```

The `-k` option assumes the user has accepted the default key file names for the SSH protocol, version 2. These file names are `id_dsa.pub` or `id_rsa.pub`, and are located in the `~/.ssh` directory.

You may be prompted to acknowledge cell authenticity, and may be prompted for the remote user password. The `-k` key exchange is done serially over the cells to prevent the user from getting password prompts from all cells simultaneously. After the `-k` option is used once, subsequent commands to the same cells do not require the `-k` option and do not require passwords for that user from the host.
9.3.2 Using dcli with the -n Option

This example shows how to run the CellCLI command `ALTER IORMPLAN OBJECTIVE='basic'`, and abbreviates non-error output.

Example 9-2 Using the -n Option

```bash
$ ./dcli -g mycells -l celladmin -n "cellcli -e alter iormplan \ objective='basic'"
```

The abbreviated output would be similar to the following:

```
OK: ['abcd2x3']
stsd2s2:
stsd2s2: CELL-02619: Current IORMPLAN state is not 'active'.
```

9.3.3 Using dcli with the -r Option

This example shows how to run the CellCLI command `LIST GRIDDISK`, and deletes the lines in the output that contain `normal`.

The command is run on the target cells listed in the `mycells` group file.

Example 9-3 Using the -r Option

```bash
$ ./dcli -l celladmin -r '^normal' -g mycells "cellcli -e list celldisk"
```

The output would be similar to the following:

```
*.normal: ['stsd2s2', 'abcd2x3']
abcd2x3: CD_06_abcd2x3    importRequired
```

9.3.4 Using dcli with the -v Option

This example shows how to use the verbose (-v) option with SSH.

Example 9-4 Using the -v Option

```bash
$ ./dcli -s "-v" -c mycell date
```

9.3.5 Using dcli with the -t Option

This example shows how to use the -t option to list target cells.

The -t option should be used with -c or -g option.
9.3.6 Using dcli with the \texttt{-f} Option

This example shows how to use the \texttt{-f} option.

Example 9-6 Using the \texttt{-f} Option

In this example, a set of files is copied to the storage servers listed in the \texttt{mycells} file.

\[ \texttt{dcli -g mycells -f '*.bin'} \]

9.3.7 Using dcli with the \texttt{--vmstat} Option

This example shows how to use the \texttt{--vmstat} option of dcli.

Example 9-7 Using the \texttt{--vmstat} Option

\[ \texttt{$ ./dcli -g 123 -l sage --vmstat='-a 3 5$} \]

\begin{verbatim}
procs ----------memory---------- ---swap-- -----io----  --system-- ...
in  cs  us
abcd2x1:  2  0  22656 178512  792272  0  0  1  21
  7  2  2
abcd2x2:  0  0  452304  21432  108760  867712  0  1  178  269
  2  0  2
abcd2x3:  0  0  49252  912164  70156  49996  1  1  74  249
  2  2  1
Minimum:  0  0  21432  70156  49996  0  0  1  21
  2  0  2
Maximum:  2  0  452304  912164  178512  867712  1  1  178  269
  7  2  2
Average:  0  0  167185  318750  119142  569993  0  0  84  179
  3  1  1
\end{verbatim}

9.3.8 Using dcli with the \texttt{--hidestderr} Option

This example shows how to use the \texttt{--hidestderr} option.

This option can only be used when SSH is used for remotely run commands.

Example 9-8 Using the \texttt{--hidestderr} Option

This first command does not use the \texttt{--hidestderr} option, so the errors are shown.

\[ \texttt{$ ./dcli -l root -g cell "ls -l unknown_file; cellcli -e list cell"} \]

exam08cel01: ls: unknown_file: No such file or directory
This second command uses the --hidestderr option, so the errors are not shown.

```sh
$ ./dcli -l root -g cell --hidestderr -l root "ls -l unknown_file;\ncellcli -e list cell"
```

9.3.9 Using dcli with the --showbanner Option

This example shows how to use the --showbanner option.

The banner of the remote cell replaces ******BANNER****** shown in the example.

**Example 9-9 Using the --showbanner Option**

```sh
$ ./dcli --showbanner -l root -g cell "cellcli -e list cell"
exam08cel01: exam08cel01 online
exam08cel01: ******BANNER******
exam08cel01: exam08cel01 online
exam08cel02: ******BANNER******
exam08cel02: exam08cel02 online
```

9.3.10 Using dcli to Change an IORM Plan

This example shows a CellCLI command that changes the IORMPLAN to active on the target cells in the mycells group file.

The -t option displays the cells that are in the mycells group file.

**Example 9-10 Using dcli to Change an IORM Plan**

```sh
$ ./dcli -g mycells -l root -t "cellcli -e alter iormplan active"
```

9.3.11 Using dcli with a Script

This example shows the CellCLI commands in the reConfig.scl file on the target cells.

The target cells are contained in the mycells group file. The command is run as the default celladmin user.

**Example 9-11 Using dcli with a Script**

```sh
$ ./dcli -g mycells -x reConfig.scl
```
9.3.12 Using dcli to List Grid Disk Status

This example shows how to run a CellCLI command that lists the name and status of grid disks on the target cells.

The target cells are contained in the mycells group file. The command is run as the default celladmin user. Output lines that contain active as the status are deleted.

Example 9-12  Using dcli to List Grid Disk Status

$ ./dcli -r '.*active' -g mycells "cellcli -e list griddisk"

9.3.13 Using dcli to List Alert History Information

This example shows a CellCLI command that lists alert history name, examined by, severity on the target cells.

The target cells are contained in the mycells group file. The command is run as the default celladmin user. Output lines that contain clear for severity are deleted.

Example 9-13  Using dcli to List Alert History Information

$ ./dcli -r '.*clear' -g mycells "cellcli -e list alerthistory attributes name, examinedby, severity"

9.3.14 Using dcli to List Alert History Where examinedby is Not Set

This example shows a CellCLI command that lists alert history where examined by has not been set on the target cells.

The target cells are contained in the allcells group file. The command is run as the default celladmin user.

Example 9-14  Using dcli to List Alert History where examinedby is not Set

$ ./dcli -g allcells -l celladmin "cellcli -e list alerthistory where examinedby=\"'

9.3.15 Using dcli to List Current Metric Alert State

This example shows a CellCLI command that retrieves metric current objects for the number of MB read in large blocks on a grid disk for a group of cells.

The command lists the current metric alert state and metric value for the metric GD_IO_BY_R_LG on the target cells. The target cells are contained in the mycells group file. The command is run as the default celladmin user.

Example 9-15  Using dcli to List Current Metric Alert State

$ ./dcli -g mycells "cellcli -e list metriccurrent GD_IO_BY_R_LG \ attributes alertstate, metricvalue"
9.3.16 Using dcli to List Specific Metric Current Objects in a Group

This example shows a CellCLI command that retrieves metric current objects for the number of requests to read or write blocks on a grid disk.

The CellCLI command lists metric current objects for names that begin with GD_IO_RQ on the target cells. The target cells are contained in the mycells group file. The command is run as the default celladmin user.

Example 9-16 Using dcli to List Specific Metric Current Objects in a Group

```bash
$ ./dcli -g mycells "cellcli -e list metriccurrent where name like '%GD_IO_RQ.*'"
```

9.3.17 Using dcli to List Specific Metric Current Objects

This example shows a CellCLI command that lists metric current objects with name equal to cl_put (cell CPU utilization) on the target cells.

The target cells are contained in the mycells group file. The command is run as the default celladmin user.

Example 9-17 Using dcli to List Specific Metric Current Objects

```bash
$ ./dcli -g mycells "cellcli -e list metriccurrent cl_cput"
```

9.3.18 Using dcli to List Physical Disks

This example shows a CellCLI command that lists physical disks where status is not equal to normal on the target cells.

The target cells are contained in the mycells group file. The command is run as the default celladmin user.

Example 9-18 Using dcli to List Physical Disks

```bash
$ ./dcli -g allcells "cellcli -e list physicaldisk where status not = normal"
```

9.3.19 Using dcli to List Cell Disks with Free Space

This example shows a CellCLI command that lists cell disks where free space is less than 100 MB on the target cells.

The target cells are contained in the mycells group file. The command is run as the default celladmin user.
Example 9-19 Using dcli to List Cell Disks with Free Space

In this example, the backslash (\) is an escape character that allows the greater than character (>) to be passed to the CellCLI utility without being interpreted by the remote shell.

$ ./dcli -g allcells "cellcli -e list celldisk where freespace > 100M"

9.3.20 Using dcli to View Alert History

This example shows a CellCLI command to view the alert history from a particular period.

Example 9-20 Using dcli to View Alert History

In this example, the backslash (\) is an escape character that allows the greater than character (>) and the quotation marks to be passed to the CellCLI utility without being interpreted by the remote shell.

dcli -g lab.cells "cellcli -e list alerthistory where begintime > \\
'Aug 4, 2009 12:06:38 PM'"

9.4 Setting Up SSH User-Equivalence on Oracle Exadata Storage Server

Setting user-equivalence enables you to issue commands to remote cells without having to enter the password for the cell.

- To set up SSH user-equivalence for use with the dcli utility, use the -k option.

See Using dcli to Set up SSH User-equivalence for a Current User.
10 Setting up Oracle Exadata Storage Snapshots

Exadata Snapshots are ideal for creating space-efficient read-only or read-write clones of an Oracle database that you can use for development, testing, or other non-production purposes.

- **Overview of Exadata Snapshots**
  Before creating Exadata snapshot databases, check that your environment meets these requirements.

- **Exadata Snapshot Concepts**
  There are various objects you need when creating and using Exadata Snapshots.

- **Sparse Disk Sizing and Allocation Methodology**
  To create Exadata Snapshots, you must have sparse grid disks created with an Oracle ASM disk group created based on those disks.

- **Refresh Considerations, or Lifecycle of Exadata Snapshots**
  Refresh cycles can influence how you use and create Exadata Snapshots.

- **Using an Oracle Data Guard Standby Database as the Test Master**
  If the test master is a complete database that needs to be refreshed regularly, Oracle recommends creating the test master database as an Oracle Data Guard physical standby dedicated to this purpose.

- **Managing Exadata Snapshots**
  To create and manage Exadata Snapshots, you need to perform these procedures:

  - **Managing Sparse Griddisks**
    You can resize, recreate, or monitor the activity of sparse griddisks.

  - **Monitoring Exadata Snapshots Using Database Statistics and Wait Events**

10.1 Overview of Exadata Snapshots

Traditionally, to clone databases in a production system, you would create test master and snapshot databases on a non-Exadata system (Figure 10-1). In some cases, these databases are a full copy that consumes as much storage as its source (Figure 10-2).
If the clones are a full copy of the production database, this is expensive in terms of the amount of storage consumed and the time it takes to create the clones.
Imagine creating ten clones for a multi-terabyte database and it is easy to see why this approach does not scale.

Another drawback to this approach is that Oracle Exadata System Software features such as Smart Scan, Smart Logging, and Smart Flash are not available on non-Exadata systems.

To solve these problems, you can use Exadata Snapshots. Exadata Snapshots are ideal for creating space-efficient read-only or read-write clones of an Oracle database that you can use for development, testing, or other non-production purposes, and when multiple clones are required because of disk space and time savings. The following image depicts the space required for an Exadata Snapshot.

**Note:**

Exadata Snapshots should be used only for development and testing purposes. They should not be used in production environments.

---

**Figure 10-3  An Exadata Snapshot**

An Exadata Snapshot is based on a test master, which is a full clone of the source database. The test master is the only full copy of the source database. From a single test master you can create multiple Exadata Snapshots with minimal additional storage and minimal effort. Each Exadata Snapshot uses a small fraction of the disk space required for the test master and can be created or dropped in seconds. Each Exadata Snapshot is a logical copy of the test master.
Before creating Exadata Snapshots from the test master, you can modify the data in the test master, if required. For example, you can delete or mask sensitive data in the test master before making it available to non-privileged users.

Creating an Exadata Snapshot from the test master is as simple as recording the parent file name in the child file header, an operation that completes in seconds and requires minimal disk space. Additional disk space is consumed only when the user of the snapshot begins to change data. Only new data is written to data blocks that are allocated to the snapshot on write. All requests for data that has not changed are serviced by the data blocks of the test master.

Multiple users can create independent snapshots from the same test master. This enables multiple development and test environments to share space while maintaining independent databases for each user. The following image shows an Exadata environment with three Exadata Snapshots that use the same test master.

**Figure 10-4  An Exadata Environment with 3 Exadata Snapshots from the Same Test Master**

![Diagram](image)

**Hierarchical and Read-Write Snapshots**

Oracle Exadata System Software release 12.2.1.1.0 introduces hierarchical and read-write snapshots. Hierarchical snapshots enable you to create snapshots from snapshots. You might want to do this if you are working on your snapshot and wish to save a copy before you make additional changes to it. In hierarchical snapshots, you can make a snapshot at any level descended from the test master. Exadata Snapshots are writable. A snapshot points to the parent's blocks for the data. If you edit a snapshot, then the snapshot will point to the new data. For the unchanged data, it will point to the parent's blocks.

If you have taken a snapshot from a snapshot, and you edit the parent snapshot, then you have to delete all snapshots that are dependent on that snapshot.
• **Exadata Snapshot Support of Exadata Features**
  In addition to space and time savings, Exadata Snapshots provide cost-effective development, quality assurance and test environments on Oracle Exadata Database Machine.

• **Separate Test/Development and Production Environments**
  Oracle recommends that test and development environments be hosted on a separate physical Oracle Exadata Rack from the rack hosting the production database.

• **Types of Exadata Snapshots**
  You can create two types of Exadata Snapshots, depending on the current setup of your environment.

• **Hierarchical Snapshot Databases**
  Hierarchical snapshots enable you to create snapshot databases from other snapshot databases.

• **Sparse Test Masters**
  With the introduction of hierarchical snapshots, you can now create sparse test masters.

### 10.1.1 Exadata Snapshot Support of Exadata Features

In addition to space and time savings, Exadata Snapshots provide cost-effective development, quality assurance and test environments on Oracle Exadata Database Machine.

Exadata Snapshots can be used by developers and testers who need to validate functionality. You can also use Exadata Snapshot to practice maintenance and operational steps in a fully functional Exadata environment (for example, Exadata Smart Flash Cache, Exadata Smart Scan Offload, and Exadata Hybrid Columnar Compression).

### 10.1.2 Separate Test/Development and Production Environments

Oracle recommends that test and development environments be hosted on a separate physical Oracle Exadata Rack from the rack hosting the production database.

An Oracle Exadata Database Machine system dedicated to development and test is ideal. Test masters and their associated Exadata Snapshots would be hosted on this system. Alternatively it may be an Oracle Exadata Database Machine system that hosts Oracle Data Guard standby databases for high availability, disaster recovery, or other purposes as permitted by capacity. Test masters and their snapshots may reside in either physical or virtual machines on an Oracle Exadata Database Machine system.

**Related Topics**

• **Types of Exadata Snapshots**
  You can create two types of Exadata Snapshots, depending on the current setup of your environment.

### 10.1.3 Types of Exadata Snapshots

You can create two types of Exadata Snapshots, depending on the current setup of your environment.
• **Using a test master from a pluggable database (PDB).**

You can create **Exadata Snapshot PDBs** from individual PDBs within a container database (CDB). You can clone individual PDBs to create the test master PDB, from which Exadata Snapshot PDBs are generated.

You can move Exadata Snapshot PDBs from one CDB to another in the same Exadata cluster. You can also create an Exadata Snapshot PDB in one container from a test master PDB residing in another container as long as both CDBs are in the same Exadata cluster.

The test master database and their Exadata Snapshots must be in the same Oracle Automatic Storage Management (Oracle ASM) cluster environment.

The following image shows a production CDB with three PDBs. Two of the PDBs (PDB1 and PDB2) have been cloned to create test master PDBs, which were then unplugged and plugged into another CDB on the test Oracle Exadata Database Machine. In this figure, six Exadata Snapshot PDBs have been created from the test master PDBs.

![Exadata Snapshot PDBs](image)

**Figure 10-5** Exadata Snapshot PDBs

• **Using a test master from a whole database (CDB or non-CDB)**

An **Exadata Snapshot database** can be a snapshot of a complete container database (CDB) or non-container database (non-CDB). A snapshot of a complete CDB includes all of the PDBs in the CDB.

**Note:**

To create a snapshot of a complete CDB, you must apply Patch 32233739 to the Oracle home directory that supports the Exadata Snapshot database.

The next figure depicts a full clone of the production database. The clone, which is the test master database, is hosted on a separate test/development system, and it is associated with six Exadata Snapshots. The test master database is created using either an Oracle Data Guard standby database (recommended if the test master will be refreshed on a regular basis) or Oracle Recovery Manager (RMAN).
The test master database and their Exadata Snapshots must be in the same Oracle ASM cluster environment.

Figure 10-6 Exadata Snapshot Databases

10.1.4 Hierarchical Snapshot Databases

Hierarchical snapshots enable you to create snapshot databases from other snapshot databases.

Oracle Exadata System Software release 12.2.1.1.0 introduced hierarchical snapshots. You might want to use hierarchical snapshots if you are working on your snapshot database and you want to save a copy before you make additional changes to it.

There is no set limit to the number of levels allowed in the hierarchy, but for performance and management reasons, a practical limit of 10 levels is recommended.
Figure 10-7  Hierarchical Snapshot Databases

A snapshot database points to its parent's blocks for data. If you make a change to a snapshot database, then the snapshot database allocates a new block for the changed data. For the unchanged data, the snapshot points to the parent's blocks. A snapshot that is several levels descended from the original test master will retrieve its data by traversing up the tree starting with the snapshot database from which it was created.
If you have taken a snapshot database from another snapshot database, and you make a change to the parent snapshot database, then you have to drop all snapshot databases that depend on that snapshot database. The parent snapshot database becomes read-only when child snapshots are created from it. If you want to write to the parent snapshot again, you must drop all child snapshots.

### 10.1.5 Sparse Test Masters

With the introduction of hierarchical snapshots, you can now create sparse test masters. Hierarchical snapshot databases provide the capability to create sparse test masters. With a sparse test master there is one full copy of data to provide the parent data for multiple levels of snapshots. When you are ready to create snapshots, you create an additional set of sparse files to receive updates from your production database via a replication technology such as Oracle Data Guard.

When you need to create a new test master of current data, rather than cloning the full production database again as of the new point in time, you mark the previous additional sparse files read only creating a sparse test master and create a new set of sparse files to be kept current. Additional snapshots on this sparse test master could then be created to be used by application developers for testing. You can repeat this process up to 9 times total to allow one full copy of the production database to support snapshots from multiple points in time. Creating the new sparse test master is very quick (5 minutes or less) allowing you to easily keep current with production.
The sparse test master database can be a container database (CDB) or non-container database (non-CDB).

**Figure 10-9  Example configuration with Sparse Test Masters**

Related Topics

- **Creating Sparse Test Masters from a Single Full Database Copy**
  You can create multiple sparse test masters from a single copy of a full database.

### 10.2 Prerequisites for Exadata Snapshot Databases

Before creating Exadata snapshot databases, check that your environment meets these requirements.

- Storage servers must be Oracle Exadata Database Machine X3-2 or later
- **Oracle Exadata System Software 12.1.2.1.0 or later for Oracle Exadata Storage Servers and Oracle Exadata Database Servers**
  You cannot downgrade to an earlier version with sparse grid disks on a cell.
- **Oracle Grid Infrastructure software release 12.1.0.2.0 BP5 or later**
  The Oracle ASM disk group that contains the sparse Oracle ASM grid disks must have both `COMPATIBLE.RDBMS` and `COMPATIBLE.ASM` set to 12.1.0.2 or later.
  The parent disk group can be 11.2 compatible.
- For disk groups hosting test master files, `COMPATIBLE.RDBMS` must be set to 11.2.0.0.0 or later.
- **Oracle Database software release 12.1.0.2.0 BP5 or later**
  The parent database and the snapshot database must be 12.1.0.2 compatible.
- The data files for the snapshot database and the parent database must be on the same Oracle ASM cluster.
• The db_block_size database initialization parameter must be at least 4096 and be a multiple of 4096.

• If you are using hierarchical snapshot databases, sparse test master databases, the new --sparse option for the Oracle ASM cp command, or the new setsparseparent Oracle ASM command, then you need Oracle Database and Oracle Grid Infrastructure software release 12.2.0.1.0 and Oracle Exadata System Software release 12.2.1.1.0 or later.

• Exadata sparse griddisks and hence Oracle ASM Sparse diskgroups cannot be created on Exadata XT storage servers.

• To create a snapshot of a complete container database (CDB), you must apply Patch 32233739 to the Oracle home directory that supports the Exadata Snapshot database.

Related Topics
• setsparseparent Command

10.3 Exadata Snapshot Concepts

There are various object you need when creating and using Exadata Snapshots.

• Sparse Database and Sparse Files
  In a sparse database, such as an Exadata Snapshot database, its data files are sparse files.

• Sparse Grid Disks
  A sparse grid disk has a virtual size attribute as well as physical size.

• Sparse Disk Groups
  Exadata Snapshots utilize Oracle ASM sparse disk groups.

10.3.1 Sparse Database and Sparse Files

In a sparse database, such as an Exadata Snapshot database, its data files are sparse files.

A database consists of the following files:

• control files
• online redo logs
• temp files
• data files

A sparse file contains only changes made to blocks from the parent file (the parent file remains unchanged) and maintains a pointer to the parent file for access to unchanged data.

The Exadata Snapshot has its own copy of the other database files (control files, online redo logs, and temp files). These other database files are not sparse files.

Related Topics
• Overview of Exadata Snapshots
10.3.2 Sparse Grid Disks

A sparse grid disk has a virtual size attribute as well as physical size.

A sparse Oracle ASM disk group is composed of sparse grid disks.

The maximum physical size that can be allocated to sparse grid disks from a single cell disk is 4 TB. The maximum allowed virtual size is 100 TB.

You create sparse grid disks before creating a sparse Oracle ASM disk group. If you are deploying a new system or re-deploying an existing system, set aside some space from each cell disk for use by the disk group just as you would for any other disk group.

If you want to use an existing system without reimaging, see My Oracle Support note 1467056.1 for instructions on resizing existing grid disks.

Related Topics
- Creating Sparse Grid Disks
- Resizing Grid Disks in Exadata: Examples (My Oracle Support Doc ID 1467056.1)

10.3.3 Sparse Disk Groups

Exadata Snapshots utilize Oracle ASM sparse disk groups.

Sparse data files can be created only in Oracle Automatic Storage Management (Oracle ASM) sparse disk groups. The following figure shows a sparse disk group containing three Exadata Snapshots.

![Sparse Disk Group Contains Exadata Snapshots](image)

A sparse Oracle ASM disk group can store both sparse and non-sparse files. You can create full database objects in a sparse Oracle ASM disk group. For example, you can
create your test master database in your sparse Oracle ASM disk group along with the Exadata Snapshots. Note that there will be no space savings for the test master database because it will be a full copy of the source database.

A sparse disk group has the following attributes:

- `compatible.asm` must be set to 12.1.0.2 or higher.
- `compatible.rdbms` must be set to 12.1.0.2 or higher.
- `cell.sparse_dg` must be set to `allsparse`. This attribute identifies the disk group to Oracle ASM as being made up of sparse grid disks.
- `appliance.mode` must be set to `true`.
- A sparse disk group uses 16 times the size of an extent. For 4M allocation unit (AU), each extent is 64M.
- A sparse disk group uses Virtually Allocated Metadata.

For example, the following SQL command creates a sparse disk group:

```
SQL> CREATE DISKGROUP SPARSE NORMAL REDUNDANCY DISK 'o/*/SPARSE_*'
         ATTRIBUTE
                 'compatible.asm'          = '12.1.0.2',
                 'compatible.rdbms'         = '12.1.0.2',
                 'cell.smart_scan_capable'   = 'true',
                 'cell.sparse_dg'            = 'allsparse',
                 'au_size'                   = '4M';
```

### 10.4 Sparse Disk Sizing and Allocation Methodology

To create Exadata Snapshots, you must have sparse grid disks created with an Oracle ASM disk group created based on those disks.

It is not possible to directly modify an existing disk group and convert it into a sparse disk group; you must create the grid disks with the sparse attributes. If you want to use an existing disk group, then the disk group and associated grid disks must be dropped and re-created.

If you want to create a new sparse disk group and all of your space is currently allocated to existing disk groups, then you must resize one or more existing disk groups to free up space. Use the following steps to perform this process.

- **Sizing Steps for New Sparse Disk Groups**
  These steps describe how to determine the space needed for sparse disk groups and how to allocate that space.

  The maximum physical size that can be allocated to sparse grid disks from a single cell disk is 4 TB.
The maximum virtual size that can be allocated per sparse grid disk is 100 TB. However, to avoid additional performance impact Oracle does not recommend creating large virtual sizes unless it is truly required.

Remember also that a sparse disk group can house non-sparse database files, in which case the physical storage utilization matches the virtual file size.

1. Determine the amount of physical space for your snapshots. Use the formula found in Calculating the Physical Size for Grid Disks.

   Note that a good maximum starting point for physical space allocation for sparse grid disk is 15% of your total space. This can be tuned up or down as required based on utilization patterns and future needs.

2. Determine the amount of virtual space you will require for your snapshots. Use the formula found in Calculating the Virtual Size for Grid Disks.

   Note that the virtual size can be modified online by modifying the Oracle ASM disk group without making any changes to the underlying grid disks.

3. Using the number of snapshots you plan on having at any one time and how much changes you expect to make, use the process defined in Determine the Amount of Available Space to determine which existing disk group(s) you want to shrink to release space for the sparse disk group. Based on your current allocations, you may need to shrink multiple disk groups to get the space you require for the sparse disk group. Leave a minimum of 15% free space in the original disk groups to allow for rebalancing during the resize operation.

   If there is not enough free space available in your current disk groups to provide desired space for the sparse disk group:
   - Rethink your sparse usage in this environment and downsize accordingly
   - Relocate objects and databases from this environment to another with available free space
   - Add storage

4. Once you have decided which disk group to resize, follow the steps in Shrink the Oracle ASM Disks in the Donor Disk Group and Shrink the Grid Disks in the Donor Disk Group to resize the disks.

   This process allows you to resize disks while they are online and only require a single rebalance operation per disk group. Oracle ASM may run more than one rebalance operation concurrently, depending on the number of Oracle ASM instances available in the cluster.

5. Create the sparse grid disk from the space just released using the commands outlined in Creating Sparse Grid Disks.

6. Create the sparse Oracle ASM disk group using the commands outlined in Resizing the Physical Space.

7. To monitor sparse grid disk activity see Monitoring Sparse Disk Group Utilization.

8. If you find that you need to resize your sparse grid disks:
   - If more physical space is required, follow steps 1 through 4 in this section to free up space and then follow the steps in Resizing the Physical Space. The maximum physical size that can be allocated to sparse grid disks from a single cell disk is 4 TB.
If more virtual space is required, follow the steps in Resizing the Virtual Space. The maximum virtual size that can be allocated to sparse grid disks from a single cell disk is 100 TB.

10.5 Refresh Considerations, or Lifecycle of Exadata Snapshots

Refresh cycles can influence how you use and create Exadata Snapshots. The test master must remain in either READONLY state (if open) or closed while it has Exadata Snapshots associated with it. If the test master needs to be refreshed, all Exadata Snapshots dependent on the test master must be dropped and re-created.

If different groups of Exadata Snapshot users have different refresh cycle requirements, you may need to maintain multiple test masters. The following figure shows three test masters, each with its own refresh schedule.

**Figure 10-11 Three Developer Groups for PDB1, Each With a Different Refresh Cycle**

10.6 Using an Oracle Data Guard Standby Database as the Test Master

If the test master is a complete database that needs to be refreshed regularly, Oracle recommends creating the test master database as an Oracle Data Guard physical standby dedicated to this purpose.

There are multiple benefits when using this approach:

- Easier refresh
  Oracle Data Guard is a proven solution for synchronizing multiple physical replicas of a production database used for disaster recovery, read-only offload from
production, backup offload, and test. This same functionality can be used to maintain a copy of the production database dedicated to serving as a test master database that is more easily refreshed on a periodic basis. The benefit of using an Oracle Data Guard physical standby increases as the size of the test master database and the number of times it must be refreshed increases. These benefits far outweigh the small incremental effort required to create an Oracle Data Guard replica compared to simply cloning a database from an Oracle Recovery Manager (RMAN) backup.

- Minimal impact to primary

  During the time the Oracle Data Guard replica is being used as the test master database, Oracle Data Guard redo transport and apply are disabled. There is zero impact to the production database. When it is time to refresh, only the deltas generated since the Oracle Data Guard replica was converted to a test master database are taken from the production database and used to resynchronize the test master database.

**Note:**

Since transport and apply for this Oracle Data Guard replica will be stopped while it functions as a test master, it should not be used for disaster recovery or for any purpose other than as test master. If you are already using Oracle Data Guard for high availability or disaster protection, Oracle recommends creating an Oracle Data Guard replica(s) to use as test master databases for Exadata Snapshot databases.

- Easy to scrub prior to creating snapshot clones

  Oracle Data Guard makes it easy to modify the test master database before making it available to create Exadata Snapshots. For example, an Oracle Data Guard replica can be opened read-write and data can be masked or scrubbed prior to creating Exadata Snapshots. Later, when testing is complete, the test master database can be converted back into an Oracle Data Guard replica discarding any modifications made to the original copy and refreshing it using only the deltas from the production database. Note that after refreshing the Oracle Data Guard replica, you need to re-scrub the database before you can use it as a test master again.

  If you are using an RMAN backup database, and you mask or scrub the data, when you need to refresh the test master, you have to create another backup as the test master and rescrub it to make it current.

**Related Topics**

- **Refreshing the (Read-only) Test Master Database**
  To refresh a read-only test master database, it must be converted temporarily to a read-write test master.

- **Creating a Physical Standby Database**

10.7 Managing Exadata Snapshots

To create and manage Exadata Snapshots, you need to perform these procedures:

- **Creating Sparse Grid Disks**
Creating an ASM Disk Group for the Sparse Grid Disks

Setting Up the Test Master
You can create the test master using one of two methods:

Creating Snapshots
You can create an Exadata Snapshot of a pluggable database (PDB) or a complete database.

Refreshing the (Read-only) Test Master Database
To refresh a read-only test master database, it must be converted temporarily to a read-write test master.

Creating a Snapshot Database from Another Snapshot Database

Creating Sparse Test Masters from a Single Full Database Copy
You can create multiple sparse test masters from a single copy of a full database.

Creating Sparse Test Masters for PDBs
This procedure creates a hierarchical snapshot tree or sparse test masters manually for a pluggable database (PDB) in an Oracle Multitenant database.

Determining All Snapshots Associated with a Test Master
Use this query to discover all of the children associated with a test master.

Doing a Sparse Copy

10.7.1 Creating Sparse Grid Disks

When creating sparse grid disks, you need to specify the physical size and the virtual size.

• Calculating the Physical Size for Grid Disks
• Calculating the Virtual Size for Grid Disks
• Creating a Sparse Grid Disk

10.7.1.1 Calculating the Physical Size for Grid Disks

You can use the following formula to get a rough estimate of the total physical space to set aside for a sparse ASM disk group:

\[
\text{Total physical space} = \left( \sum \text{size of all test masters in the sparse ASM disk group} \right) + \left( \sum \text{approximate size of all updates to the snapshot databases} \right) \times \text{ASM Redundancy}
\]

In the formula above, ASM redundancy takes into account ASM mirroring of extents. Exadata requires ASM redundancy set to either normal redundancy (double mirror the extents) or high redundancy (triple mirror the extents). If the sparse ASM disk group will use normal redundancy, expect to double the space used. If using high redundancy, expect to triple the space used.

For example, if you want 2 test masters in the sparse ASM disk group created with normal redundancy with a combined total space of 500 GB (250 GB each) and each test master will have 5 Exadata snapshots with the expectation that each snapshot will modify 20% of the blocks, then the total physical space that will be needed can be calculated as follows:
Space for 2 test masters: 2 * 250 GB = 500 GB
Space for 5 snapshots per test master, for a total of 10 snapshots:
   10 * 250 GB * 20% = 500 GB
Subtotal
   1000 GB
Normal redundancy: 2 * 1000 GB = 2000 GB

Divide this value by the number of disks to determine the size parameter for each disk. ASM grid disks should be allocated on 16 MB boundaries. If the size parameter in MB for each grid disk is not evenly divisible by 16, adjust up to a 16 MB boundary.

Note that you should set aside additional space to use for multiple projects and through multiple iterations.

Also to accommodate any disk rebalancing operations, you should add a 15% space cushion on top of the space used for snapshots and test masters.

10.7.1.2 Calculating the Virtual Size for Grid Disks

You can use the following formula to get a rough estimate of the virtual size to assign for a sparse ASM disk group:

Virtual size required for sparse disks = 
   (SUM(full virtual size of all Exadata snapshots) + Physical space allocated) 
   * ASM Redundancy

To continue with the example from the previous section, you have 10 Exadata snapshots. If they were full copies of the test master, they would be 250 GB each.

The following shows the calculation for the total virtual space:

Full size for 5 snapshots per test master, for a total of 10 snapshots:
   10 * 250 GB = 2500 GB
Size of the 2 test masters: 2 * 250 GB = 500 GB
Subtotal
   3000 GB
Normal redundancy: 2 * 3000 GB = 6000 GB

Divide this value by the number of disks to determine the virtualSize parameter for each disk. The virtual size for each grid disk should be allocated on 16 MB boundaries. If the virtualSize parameter in MB for each grid disk is not evenly divisible by 16, adjust up to a 16 MB boundary.

Note that you should set aside additional space to use for multiple projects and through multiple iterations.

10.7.1.3 Creating a Sparse Grid Disk

To create the sparse ASM grid disk, log in to each cell (or use dcli) in CellCLI and run a command similar to the following, changing size values where appropriate:

CellCLI> create griddisk all harddisk prefix=SPARSE, size=56G, virtualsize=560G

This creates a grid disk of physical size 56 GB but presents to ASM as a 560 GB grid disk. The size parameter should match the actual physical size of the ASM grid disk while the virtualSize parameter should be at least the physical size of the ASM grid disk.

Attributes for the ASM grid disk created above would look like the following from the "LIST GRIDDISK DETAIL" command:
CellCLI> LIST GRIDDISK DETAIL
size: 56G
sparse: TRUE
virtualSize: 560G

size displays the actual physical size of the grid disk.
sparse has a value of TRUE.
virtualSize displays the virtual size of the grid disk.

10.7.2 Creating an ASM Disk Group for the Sparse Grid Disks

After you have created the sparse grid disks, you create an ASM disk group to enable those disks to be accessible to a database on the cluster. To create a disk group, log in to an ASM instance using SQL*Plus as sysasm and run a command similar to the following:

```sql
SQL> create diskgroup SPARSE high redundancy disk 'o/*/SPARSE_*' attribute
'compatible.asm'='12.1.0.2',
'compatible.rdbms'='12.1.0.2',
'au_size'='4M',
'cell.smart_scan_capable'='true',
'cell.sparse_dg'='allsparse',
'appliance.mode' = 'TRUE';
```

**compatible.asm** must be set to 12.1.0.2 or higher.

**compatible.rdbms** must be set to 12.1.0.2 or higher.

**cell.sparse_dg** must be set to "allsparse". This identifies the disk group to ASM as being made up of sparse grid disks.

**appliance.mode** must be set to true.

10.7.3 Setting Up the Test Master

You can create the test master using one of two methods:

- **Create a New Test Master - Full Clone on a Disk Group with ASM ACL Enabled**
- **Converting an Existing Full Clone or Standby Database to a Test Master**
  If you already have a full clone or a standby database that you want to repurpose as a test master, then you can convert that database to a test master.
- **Setting the Ownership of the Test Master Data Files**
  After you have cloned the database to create a test master database, configure permissions on the disk group and data files.
10.7.3.1 Create a New Test Master - Full Clone on a Disk Group with ASM ACL Enabled

You can create a full clone of your database using RMAN backup/restore, data pump, or any of the other methods typically used to create a full clone of a database.

After creating the full clone, remove write permissions on all the data files to help prevent accidental overwrite.

SQL commands in ASM instance only allow to set file permissions to read only. You cannot remove write permissions in SQL.

```
SQL> ALTER DISKGROUP DATA set permission owner=read ONLY, group=read ONLY, other=none for file 'FILENAME';
```

10.7.3.2 Converting an Existing Full Clone or Standby Database to a Test Master

If you already have a full clone or a standby database that you want to repurpose as a test master, then you can convert that database to a test master.

Standby databases cannot be running redo apply while serving as a test master.

1. If you are using an Oracle Data Guard standby database, perform the following steps:
   a. For initial creation of the Oracle Data Guard replica, use the steps outlined in My Oracle Support note 1617946.1.
      The Data Guard copy must have enough redo applied that it can be opened in a READ ONLY state.
   b. If the test master is a physical standby database and you need to make any modifications to the test master, for example, deleting or masking sensitive data, then perform the following steps:
      i. Convert the standby database into a snapshot standby.

         Note:

         An Oracle Data Guard snapshot standby is different from an Oracle Exadata snapshot. An Oracle Data Guard snapshot standby is a complete copy of the source database that is open read-write. Conversion to an Oracle Data Guard snapshot standby is a simple operation using a single command. Oracle Data Guard snapshot standby facilitates making modifications to the test master and refreshing it for subsequent rounds of testing. See Oracle Data Guard Concepts and Administration (referenced at the end of this topic) for more information on Oracle Data Guard snapshot standby databases.

      ii. Modify the standby database as required.
c. When the standby database is at a consistent state and can be opened in 
READ ONLY mode, stop log transport to the standby and disable redo apply on 
the standby.

DGMGRL> edit database TESTMASTER set property logshipping=OFF; 
Property "logshipping" updated 

If you have not converted the physical standby database into a snapshot 
standby, then stop redo apply.

DGMGRL> edit database TESTMASTER set state=APPLY-OFF; 
Succeeded 

2. If access control is not already enabled on the disk group that contains the test 
master's data files, then enable access control on the disk group.

The disk group must be on Oracle Exadata storage servers.

SQL> ALTER DISKGROUP DATA SET ATTRIBUTE 'ACCESS_CONTROL.ENABLED' = 'TRUE'; 

3. Grant ownership to all data files.

See Setting the Ownership of the Test Master Data Files for details.

4. Remove write permissions on all the data files to help prevent accidental 
overwrite.

SQL commands in an Oracle ASM instance only allow you to set file permissions 
to read only. You cannot remove write permissions in SQL.

SQL> ALTER DISKGROUP DATA set permission owner=read ONLY, 
group=read ONLY, other=none for file 'FILENAME';

This allows snapshots to be created and owned by users other than the owner of 
the base files.

Related Topics

• Creating a Physical Standby using RMAN Duplicate (RAC or Non-RAC) (My 
Oracle Support Doc ID 1617946.1)

• Managing a Snapshot Standby Database

10.7.3.3 Setting the Ownership of the Test Master Data Files

After you have cloned the database to create a test master database, configure 
permissions on the disk group and data files.

Set an operating system user as the owner of the disk group, and make the operating 
system user the owner of the test master's data files.

You can do this by running SQL commands manually in SQL*Plus, or by running a 
script:
10.7.3.3.1 Running Commands Manually

You can use SQL*Plus to manually run the commands to set the ownership of the test master data files.

The following commands are run in SQL*Plus.

1. If the operating system user you are granting access to is not added as a user on the disk group, then add the user.

   \[\text{Note:}\]
   When enabling access control, all software owners that are running databases must be added as a user to the disk group.

   For example, to add the user SCOTT as an owner of the DATA disk group, use the following command:

   \[\text{SQL> ALTER DISKGROUP DATA ADD USER 'scott';}\]

2. Make the operating system user the owner of the test master's data files:

   \[\text{SQL> ALTER DISKGROUP DATA SET OWNERSHIP OWNER='scott' FOR FILE '+DATA/TESTMASTER/DATAFILE/system.257.865863315'};\]

   \[\text{SQL> ALTER DISKGROUP DATA SET OWNERSHIP OWNER='scott' FOR FILE '+DATA/TESTMASTER/DATAFILE/sysaux.258.865863317'};\]

   \[\text{SQL> ALTER DISKGROUP DATA SET OWNERSHIP OWNER='scott' FOR FILE '+DATA/TESTMASTER/DATAFILE/sysext.259.865863317'};\]

   \[\text{SQL> ALTER DISKGROUP DATA SET OWNERSHIP OWNER='scott' FOR FILE '+DATA/TESTMASTER/DATAFILE/tbs_1.256.865863315'};\]

   \[\text{Related Topics}\]
   - Oracle Automatic Storage Management Administrator's Guide

10.7.3.3.2 Running from a Script

You can also set the ownership of the test master data files using a SQL script.

The following procedure is equivalent to the commands in the previous topic, but it queries V$DATAFILE for the filenames:
1. Add an operating system user as owner of the disk group.

   SQL> ALTER DISKGROUP DATA ADD USER 'scott';

2. Generate a script called set_owner.sql to set the owner of the test master's data files.

   • If the test master is a full database, run the following in the test master database:

     ```sql
     set newpage 0
     set linesize 999
     set pagesize 0
     set feedback off
     set heading off
     set echo off
     set space 0
     set tab off
     set trimspool on
     spool set_owner.sql
     select 'ALTER DISKGROUP DATA set ownership owner='||'||||' for file '||'||||' name||';'
     from v$datafile;
     exit
     ```

   • If the test master is a PDB, run the following in the CDB$ROOT of the test master PDB:

     In the `select` statement below, the example assumes the test master PDB has a `con_id` of 10.

     ```sql
     set newpage 0
     set linesize 999
     set pagesize 0
     set feedback off
     set heading off
     set echo off
     set space 0
     set tab off
     set trimspool on
     spool set_owner.sql
     select 'ALTER DISKGROUP DATA set ownership owner='||'||||' for file '||'||||' name||';'
     from v$datafile where con_id=10;
     exit
     ```

3. Remove extra lines in `set_owner.sql`.

   ```bash
   sed -i '/SQL/d' set_owner.sql
   ```

4. Run the script in the ASM instance.

   ```sql
   SQL> @set_owner
   ```
10.7.4 Creating Snapshots

You can create an Exadata Snapshot of a pluggable database (PDB) or a complete database.

- Creating a Snapshot of a Pluggable Database
- Creating a Snapshot of a Full Database

You create an Exadata snapshot of a full database.

10.7.4.1 Creating a Snapshot of a Pluggable Database

Creating an Exadata snapshot of a pluggable database (PDB) is the simplest method for creating a snapshot because it requires no additional manual steps. Two new clauses to the `CREATE PLUGGABLE DATABASE` statement were added to identify the PDB as an Exadata snapshot. The snapshot creation process alters the permissions on the files on the test master PDB to `READONLY` to ensure that it is not modified.

Creating individual Exadata snapshot PDBs is best used when creating snapshots for a smaller number of PDBs within a given CDB. The following figure shows a high-level example of a typical lifecycle for a PDB with two PDB snapshots.

**Figure 10-12   Lifecycle of Exadata Snapshots Using PDBs**

One of the benefits of Oracle Multitenant and PDBs is the ability to easily clone an existing PDB to create a test master and move it from one CDB to another. Oracle recommends that you clone your source PDB to create the test master and then migrate it to your test environment where you can perform any data scrubbing that may be needed. Once complete, the test master PDB can be used to create numerous Exadata snapshot PDBs.
An Exadata snapshot PDB is created as a PDB in the CDB. Like all other PDBs, an Exadata snapshot PDB is subject to the general limit for the maximum number of PDBs in a single CDB. For Oracle Database 12c release 1 (12.1), the limit is 252 PDBs in a single CDB. For Oracle Database 12c release 2 (12.2) and later, the limit is 4096 PDBs in a single CDB.

After creating the test master PDB, perform the following steps to create an Exadata snapshot PDB:

1. In SQL*Plus, connect to the CDB root container (cdb$root).

   SQL> alter pluggable database PDB1TM1 close instances=all;

2. Close the test master PDB in all instances.

3. Open the test master PDB in the local instance in read only mode.

   SQL> alter pluggable database PDB1TM1 open read only;

4. Create an Exadata snapshot PDB of the test master.

   SQL> create pluggable database PDB1S1 from PDB1TM1 tempfile reuse create_file_dest='+SPARSE' snapshot copy;

   The create_file_dest argument must specify the name of a sparse disk group. The snapshot copy clause creates the PDB as a snapshot rather than a full PDB clone.

   When an Exadata snapshot PDB is created, the command changes file permissions on the test master PDB, marking the files as READONLY in ASM.

   If you need to create more PDBs than can fit in a single CDB, you can create the Exadata snapshot PDB in another CDB by creating a remote clone using a database link. For example:

   SQL> create pluggable database PDB1S1 from PDB1TM1@cdb1_dblink tempfile reuse create_file_dest='+SPARSE' snapshot copy;

### 10.7.4.2 Creating a Snapshot of a Full Database

You create an Exadata snapshot of a full database.

The following figure shows the lifecycle of an Exadata test master database and snapshot databases where the test master is based on an Oracle Data Guard replica.
Note that the test master database cannot be used as a failover target to provide high availability or disaster recovery (a Data Guard configuration may have multiple replicas that can each serve different purposes). Similar to test master PDBs, test master databases cannot be modified while Exadata snapshots exist against them.

The test master database cannot be a read-only physical standby database that is in recovery mode (for example, Active Data Guard in Real Time Apply).

The test master database and their Exadata snapshots must be in the same Oracle ASM cluster environment.

---

**Note:**

If the test master is an Oracle RAC database, you have to do one of the following:

- Create redo logs for all threads in the `CREATE CONTROLFILE` statement for the sparse clone, OR
- Specify Oracle Managed Files for the redo logs using the `ONLINE_LOG_CREATE_DEST_1` initialization parameter in the SPFILE of the sparse clone to have the redo logs created automatically.

1. In the test master database, create a sample control file script to use for your Exadata snapshot databases by backing up the existing control file to trace.

   Connect to the test master database via SQL*Plus as SYSDBA and do the following:
a. Determine name and location of any trace file to be generated by your session.
   For example:

   SQL> SELECT value FROM v$diag_info WHERE name = 'Default Trace File';

   VALUE
   /u01/app/oracle/diag/rdbms/TESTMASTER/TESTMASTER1/trace/TESTMASTER1_ora_26756.trc

b. Run the BACKUP CONTROLFILE TO TRACE command to place the CREATE CONTROLFILE command into the trace file.

   SQL> ALTER DATABASE BACKUP CONTROLFILE TO TRACE;

c. Retrieve the file shown for the Default Trace File.

2. In the test master database, determine the existing file names for the rename that will happen in step 11.
   For example, log into SQL*Plus as SYSDBA and run the following:

   SET newpage 0
   SET linesize 999
   SET pagesize 0
   SET feedback off
   SET heading off
   SET echo off
   SET space 0
   SET tab off
   SET trimspool on
   SPOOL rename_files.sql
   SELECT 'EXECUTE dbms_dnfs.clonedb_renamefile -
   ('||'|'.'||'|''||'|''|'
   REPLACE(REPLACE(REPLACE(REPLACE(name,'.','_'),-
   'TESTMASTER','SNAPTEST'),'+DATA','+SPARSE')||'|'')');'
   FROM v$datafile;
   EXIT

   The example script builds a file named rename_files.sql that contains statements for each data file similar to the following:

   EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TESTMASTER/DATAFILE/system.257.865863315',
   '+SPARSE/SNAPTEST/DATAFILE/system_257_865863315');

   In the example script, the REPLACE function:
   • Replaces periods with underscores in the original file name
   • Replaces the original database name of TESTMASTER with SNAPTEST
• Replaces the original disk group name of +DATA with +SPARSE

If you modify the example to suite your environment, ensure that you maintain consistency throughout the rest of the procedure.

3. Create an init.ora file with the contents of the test master SPFILE.

   SQL> CREATE PFILE = 'init_TestMaster.ora' FROM SPFILE;

4. Shut down the test master.

   SQL> shutdown;

5. Create an init.ora file for the Exadata snapshot database.

   You can use the init.ora file of the test master as a template, but make sure to change the db_name and control_files entries.

   For this procedure, the init.ora file for the Exadata snapshot database is referenced as snap_init.ora in commands and examples.

   $ cp init_TestMaster.ora snap_init.ora

   Modify snap_init.ora with the new database name, new control file name, and audit file destination, for example:

   db_name = SNAPTEST
   control_files = '+DATA/SNAPTEST/control1.f'
   audit_file_dest=/u01/app/oracle/admin/snaptest/adump

   You can save the modified snap_init.ora file as a template for creating additional snapshot copies of the test master.

6. Create a script to create a control file for the Exadata snapshot database.

   You will use this script later in step 10.

   a. Examine the trace file generated in step 1 and locate the CREATE CONTROLFILE command that includes the RESETLOGS clause.

      The required command is located in the trace file immediately after the following comments:

      --     Set #2. RESETLOGS case
      --
      -- The following commands will create a new control file and use it
      -- to open the database.
      -- Data used by Recovery Manager will be lost.
      -- The contents of online logs will be lost and all backups will
      -- be invalidated. Use this only if online logs are damaged.

   b. Copy the CREATE CONTROLFILE command into a new script file.

      Copy only the complete CREATE CONTROLFILE command and discard all surrounding comments and commands.

      For example, the script file may be named crt_ctlfile.sql.
c. Modify the `CREATE CONTROLFILE` command to create a control file for the Exadata snapshot database.

The control file should be created with the Exadata snapshot database name, new log file names, and the data file names of the test master. The new log files can be in any disk group that has enough space, but they should not be created in the sparse Oracle ASM disk group.

The following shows an example `CREATE CONTROLFILE` command. In the example, `SNAPTEST` is the Exadata snapshot database. The `LOGFILE` lines specify the new log file locations, and the `DATAFILE` lines specify the data file locations for the test master database.

```
CREATE CONTROLFILE REUSE SET DATABASE SNAPTEST RESETLOGS
ARCHIVELOG
  MAXLOGFILES 32
  MAXLOGMEMBERS 2
  MAXINSTANCES 1
  MAXLOGHISTORY 908
LOGFILE
  GROUP 1 '+DATA/SNAPTEST/t_log1.f' SIZE 100M BLOCKSIZE 512,
      GROUP 2 '+DATA/SNAPTEST/t_log2.f' SIZE 100M BLOCKSIZE 512
DATAFILE
  '+DATA/TESTMASTER/DATAFILE/system.257.865863315',
  '+DATA/TESTMASTER/DATAFILE/sysaux.258.865863317',
  '+DATA/TESTMASTER/DATAFILE/sysext.259.865863317',
  '+DATA/TESTMASTER/DATAFILE/tbs_1.256.865863315'
CHARACTER SET WE8DEC;
```

You can save the modified script file as a template for creating additional snapshot copies of the test master.

7. Create the `audit_file_dest` directory on all nodes on which the snapshot will be running.

   `$ mkdir -p /u01/app/oracle/admin/snaptest/adump`

8. Create the directories in Oracle ASM for the snapshot data files.

   For example:

   ```
   $ asmcmd -p
   ASMCMD > cd +SPARSE
   ASMCMD [+sparse] > mkdir SNAPTEST
   ASMCMD [+sparse/snaptest] > mkdir DATAFILE
   ```

   If the test master is a container database (CDB), create additional data file sub-directories for each pluggable database (PDB). Use the globally unique identifier (GUID) for each PDB, which is available in the `V$PDBS` view.

   For example:

   ```
   $ sqlplus / as sysdba
   SQL> ALTER SESSION set container=PDB1;
   SQL> SELECT guid FROM v$pdbs;
   ```
9. Start a database instance pointing to the Exadata snapshot database `init.ora` file (snap_init.ora) using the following commands:

```
$ sqlplus / as sysdba
SQL> startup nomount pfile=snap_init.ora
```

10. Create the Exadata snapshot control file using the script created in step 6.

    In the following example the script is named `crt_ctlfile.sql`.

    ```
    SQL> @crt_ctlfile
    ```

11. Run the script you modified in step 2.

    All the files must be renamed prior to opening the Exadata snapshot database.

    Connect using SQL*Plus as SYSDBA to the Exadata snapshot database and run the following command:

    ```
    SQL> @rename_files
    ```

    This script modifies the permissions of the test master database files stored in Oracle ASM, marking them as READONLY.

    The `dbms_dnfs.clonedb_renamefile` procedure, which is called by `rename_files.sql`, sets up the parent-child relationship between the test master database and the snapshot database, and renames the file names in the snapshot database's control file.

12. Open the Exadata snapshot database with the RESETLOGS option:

    ```
    SQL> ALTER DATABASE OPEN RESETLOGS;
    ```

13. Confirm that the Exadata snapshot files are child files of the test master database.

    Connect using SQL*Plus as SYSDBA to the Exadata snapshot, and run the following command:

    ```
    SQL> SELECT filenumber num, clonefilename child, snapshotfilename parent FROM V$CLONEDFILE;
    ```
The following is an example of the output from the query:

```
NUM CHILD
---- ---------------------------------------------
PARENT
-----------------
1   +SPARSE/SNAPTEST/DATAFILE/system_257_865863315
    +DATA/TESTMASTER/DATAFILE/system.257.865863315
2   +SPARSE/SNAPTEST/DATAFILE/sysaux_258_865863317
    +DATA/TESTMASTER/DATAFILE/sysaux.258.865863317
3   +SPARSE/SNAPTEST/DATAFILE/sysext_259_865863317
    +DATA/TESTMASTER/DATAFILE/sysext.259.865863317
4   +SPARSE/SNAPTEST/DATAFILE/tbs_1_256_865863315
    +DATA/TESTMASTER/DATAFILE/tbs_1.256.865863315
```

14. Log in using SQL*Plus to the Exadata snapshot database, and add temp files to the TEMP tablespace. This is a full size temp file, not a sparse temp file.

```
SQL> ALTER TABLESPACE temp ADD TEMPFILE '+DATA' SIZE 10G;
```

Additionally, for a snapshot of a complete CDB, connect to each PDB and add a temp file to the PDB-specific TEMP tablespace.

For example:

```
SQL> ALTER SESSION set container=PDB1;
SQL> ALTER TABLESPACE temp ADD TEMPFILE '+DATA' SIZE 10G;
```

10.7.5 Refreshing the (Read-only) Test Master Database

To refresh a read-only test master database, it must be converted temporarily to a read-write test master.

To refresh the (read-only) test master database, the main steps are:

- **Drop the Snapshot Databases**
  Delete the Exadata snapshot databases that are children of the test master database you want to refresh.

- **Change the Permissions on the Test Master to Read-Write**
  To modify the data files from read-only to read-write you can use SQL to generate a script of SQL commands.

- **Convert the Test Master Database Back to a Data Guard Replica**
  Now that the data files for the test master are writable, convert the read-only test master database into an Oracle Data Guard replica.

- **Update the Test Master Database**

- **Close the Test Master and Make All Test Master Data Files Read-Only**
  After the Test Master has been updated, you can revert it to a read-only test master.
10.7.5.1 Drop the Snapshot Databases

Delete the Exadata snapshot databases that are children of the test master database you want to refresh.

You can delete the snapshot databases using RMAN.

1. Connect to the Exadata snapshot database using RMAN with the Exadata snapshot as the target.

   Connect as a user that has the necessary privileges to start and drop a database, such as SYSBACKUP or SYSDBA.

   RMAN> CONNECT TARGET "user@snapdb_name AS SYSDBA"

2. Start the snapshot database in restricted mode.

   RMAN> STARTUP FORCE MOUNT DBA

3. Delete the snapshot database.

   RMAN> DROP DATABASE;

Note:

- Failure to drop an Exadata snapshot database has no impact on the state of a test master database. However, an Exadata snapshot can behave unpredictably if its test master database is dropped or refreshed.

- If you have a snapshot hierarchy, you must delete all of the snapshots at all levels of the snapshot hierarchy that are children of the test master database you want to refresh.

10.7.5.2 Change the Permissions on the Test Master to Read-Write

To modify the data files from read-only to read-write you can use SQL to generate a script of SQL commands.

Before starting this part of refreshing the read-only test master database, you must first drop the snapshot databases.

1. After all Exadata snapshot databases have been deleted, start up the test master database in mount mode.

2. Create a script to reset permissions on the data files for the test master database.
Connect to the test master using SQL*Plus and run the following script to create a new SQL script that contains the commands to reset permissions on the data files belonging to the test master database.

```sql
set newpage 0
set linesize 999
set pagesize 0
set feedback off
set echo off
set space 0
set tab off
set trimspool on
spool change_perm.sql
select ©ALTER DISKGROUP DATA set permission owner=read write,
group=read write, other=none for file '||'||name||'||'; from
v$datafile;
exit
```

3. Remove the extra lines from the generated script.

Run the following `sed` command to remove extra lines from the `change_perm.sql` script that you created in the previous step. Run this command from an operating system prompt, in the directory that contains the `change_perm.sql` script.

```
sed -i '/SQL/d' change_perm.sql
```

4. Use the generated script to change the file permissions.

Use SQL*Plus to connect to an Oracle ASM instance as a SYSASM user. Run the `change_perm.sql` script. This script changes the permissions of the test master's data files to make them writable.

```
SQL> @change_perm
```

10.7.5.3 Convert the Test Master Database Back to a Data Guard Replica

Now that the data files for the test master are writable, convert the read-only test master database into an Oracle Data Guard replica.

If you had originally prepared the test master database using Oracle Data Guard snapshot standby, then convert it back to its original state as an Oracle Data Guard replica using the `CONVERT` command. This command discards any changes previously made to the replica to prepare it to be the test master. It also makes it possible to refresh the test master using just incremental changes from the source database instead of a complete restore from a current backup.

10.7.5.4 Update the Test Master Database

You have two options for refreshing the test master database:

- **Allow Oracle Data Guard to refresh the test master database**

  If the Oracle Data Guard replica has been used as a test master database for only a short period of time and you have all the redo generated during this time in archive logs on disk at the source database, then you can enable redo shipping
and start redo apply. The test master database will use regular Oracle Data Guard 
protocols to retrieve archive logs and apply the logs until it is caught up with the 
primary database. Once the Oracle Data Guard replica is as current as you need 
it to be, disable redo shipping, stop redo apply and repeat the test master and 
snapshot creation cycle described in Setting Up the Test Master and Creating 
Snapshots.

This option has the benefit of being able to stop redo apply at some intermediate 
point rather than bringing the test master database totally current.

To let Oracle Data Guard refresh the standby, enable log shipping to the standby 
and redo apply on the standby:

```
DGMGRL> edit database TESTMASTER set property logshipping=ON;
Property "logshipping" updated
DGMGRL> edit database TESTMASTER set state=apply-on;
Succeeded
```

- **Use RMAN RECOVER...FROM SERVICE to roll forward the test master database**

If the Oracle Data Guard replica has been used as a test master database for a 
long period of time or if you no longer have the redo available on disk to enable 
Oracle Data Guard to automatically refresh the test master database, use RMAN 
to perform live incremental apply over the network.

A major advantage to using this method is that no additional disk space is 
required. RMAN will bring changed blocks to the standby from the primary over 
the network and apply them directly. Also RMAN greatly simplifies the process by 
determining which blocks need to be retrieved based on the SCN of the data files 
on the test master. With this method you cannot recover to an intermediate point 
in time; the refresh will bring the test master database current with the primary. For 
more information on this method refer to Performing RMAN Recovery: Advanced 

To refresh the test master database using RMAN Network Incrementals:

1. **Prepare Oracle Net Services for the RMAN connections.**

   These steps need to be performed only once.

   a. **Create a listener.ora entry for the test master database (the Oracle Data 
      Guard replica).**

      The listener entry allows RMAN to connect to the target using the SID 
      because the service is not started when the database is opened in NOMOUNT 
      mode. The following is an example of the entry:

      ```
      SID_LIST_LISTENER =
      (SID_LIST =
       (SID_DESC =
        (SID_NAME = TESTMASTER1)
        (ORACLE_HOME = /u01/app/oracle/product/12.1.0.2/dbhome_1)
       )
      )
      ```

   b. **Reload the listener to pick up the changes to the listener.ora.**

      ```
      $ lsnrctl reload listener
      ```
c. Create a TNS entry on the test master environment pointing to the SID of the local test master instance.

The entry should use the local host name rather than the SCAN name to ensure the connection request goes to the correct host.

```sql
TESTMASTER1 =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = standbydb01.example.com)
      (PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SID = TESTMASTER1)
      (UR=A)
    )
  )
```

2. Connect via RMAN to the test master database and save the CURRENT_SCN for later.

This value will be used to determine if newly created files since the last refresh need to be restored from the source database.

```sql
RMAN> select current_scn from v$database;
CURRENT_SCN#  
------------------
17081990
```

3. List the names and group identifiers of the redo log files.

The names of the online redo log files and standby redo log files of the Oracle Data Guard replica might be required in a later step.

```sql
RMAN> SELECT type, group#, member FROM v$logfile;
```

4. Refresh the standby control file of the Oracle Data Guard replica from the source database to make the control file current.

a. Reconnect to the Oracle Data Guard replica as the RMAN target.

```sql
RMAN> startup nomount force;
```

b. Restart the target in NOMOUNT mode.

```sql
RMAN> RESTORE STANDBY CONTROLFILE FROM SERVICE SOURCEMASTER;
```

c. Mount the Oracle Data Guard replica.

```sql
RMAN> ALTER DATABASE MOUNT;
```
5. Update the names of the data files and the temp files in the standby control file.

If you are not using an RMAN catalog, the names of files in the standby control file are the names that were used in the source database, not the standby.

Use the CATALOG command and the SWITCH command to update all the data file names. The SWITCH command will be used after restoring any newly created files from the source database in step 7.

In the following example, +DATA/TESTMASTER/DATAFILE/ is the location of the data files on the Oracle Data Guard replica. All data files must be stored in this location.

RMAN> CATALOG START WITH '+DATA/TESTMASTER/DATAFILE/';

6. Determine if new files were added that need to be restored from the source database.

Use the CURRENT_SCN from step 2.

RMAN> SELECT file# FROM v$datafile WHERE creation_change# >= 17081990;
FILE#
-------
  9
 10

7. If there are files returned by the previous query, restore those data files from the source database.

Execute an RMAN RUN block similar to the following using the list of FILE# values returned by the previous step. If no FILE# values were returned, then skip this step.

RMAN> run{
  2> set newname for database to '+DATA';
  3> restore datafile 9,10 from service SOURCEMASTER;
  4> }

8. If not using an RMAN catalog, rename the data files in the standby control file.

Switch to the copies cataloged in step 5.

RMAN> SWITCH DATABASE TO COPY;

9. Update the names of the online redo logs and standby redo logs in the standby control file.

Use one of the following methods:

- Use the ALTER DATABASE CLEAR command to clear the log files in all redo log groups of the Oracle Data Guard replica. RMAN then recreates all the standby redo logs and the online redo log files.
Note:

Clearing log files is recommended only if the Oracle Data Guard replica does not have access to the online redo log files and standby redo log files of the source database. If the Oracle Data Guard replica has access to the redo log files of the source database and the redo log file names of the source database are OMF names, then the `ALTER DATABASE` command will delete log files on the source database.

Also, the clearing of the log files will create new log files. Any existing log files are not used because the control file is not aware of those existing files. To conserve space, delete the existing log files from Oracle ASM prior to running the `ALTER DATABASE CLEAR` commands.

The `GROUP#` column of the `V$LOGFILE` view queried in step 5 provides the redo log group identifiers of the log groups that must be cleared. Use separate `ALTER DATABASE CLEAR` commands to clear each redo log group.

For example, the following command clears the redo log group with identifier 2.

```sql
SQL> ALTER DATABASE CLEAR LOGFILE GROUP 2;
```

- Use the `ALTER DATABASE RENAME FILE` command to rename the redo log files. Use a separate command to rename each log file listed in step 5.

To rename log files, the `STANDBY_FILE_MANAGEMENT` initialization parameter must be set to `MANUAL`. Renaming log files is recommended when the number of online redo logs files and standby redo log files is the same in the source database and the Oracle Data Guard replica.

10. Use RMAN `RECOVER....FROM SERVICE` to roll forward the data files to current state.

No additional space is required for this operation. Note that this process can only bring the files totally current; it cannot bring the files to a previous point in time. Connect via RMAN to the Oracle Data Guard replica as target using the TNS entry created in step 3. The service specified should point to the primary.

```sql
RMAN> recover database noredo from service SOURCEMASTER;
```

11. Enable redo shipping to the Oracle Data Guard replica and start redo apply. This is necessary to update the control file with the blocks applied as part of step 10.

```sql
DGMGRL> edit database TESTMASTER set property logshipping=ON;
Property "logshipping" updated
DGMGRL> edit database TESTMASTER set state=apply-on;
Succeeded.
```
12. After redo has been applied, repeat the process you used to convert the Oracle Data Guard replica into a test master database and then create Exadata database snapshots. Remember to once again disable log shipping and redo apply at the standby.

10.7.5.5 Close the Test Master and Make All Test Master Data Files Read-Only

After the Test Master has been updated, you can revert it to a read-only test master. Complete one of the tasks described in "Setting the Ownership of the Test Master Data Files".

10.7.5.6 Re-create All Snapshots

After the Test Master has been updated and made read-only again, re-create all the snapshot databases to get the latest updates.

You can create an Exadata snapshot database of a pluggable database (PDB) or of a full database as described in "Creating Snapshots."

10.7.6 Creating a Snapshot Database from Another Snapshot Database

To create a snapshot from a snapshot:

1. Create a first level snapshot. In the following example, the snapshot is called PDB1S1.

   create pluggable database PDB1S1
   from PDB1TM1
   create_file_dest='+SPARSE'
   snapshot copy;

2. Open and close the PDB, so you can re-open it as read-only in the next step.

   alter pluggable database PDB1S1 open;
   alter pluggable database PDB1S1 close;

3. Open the PDB in read-only mode so it can serve as a test master.

   alter pluggable database PDB1S1 open read only;

4. Create a snapshot from the snapshot created in step 1. In the following example, the second level snapshot is called PDB1S1_A.

   create pluggable database PDB1S1_A
   from PDB1S1
   create_file_dest='+SPARSE'
   snapshot copy;

   alter pluggable database PDB1S1_A open;

Related Topics

- Hierarchical Snapshot Databases
  Hierarchical snapshots enable you to create snapshot databases from other snapshot databases.
10.7.7 Creating Sparse Test Masters from a Single Full Database Copy

You can create multiple sparse test masters from a single copy of a full database. In this procedure, the source for the test masters is a full copy of a Data Guard physical standby database. This standby database should not be used as a target for switchover or failover; it should only be used as the source for the test masters defined in this process.

This process takes advantage of the hierarchical snapshot functionality to allow redo to be shipped and applied, keeping the standby database current with production while also providing files to be used as source for test masters to be used by Exadata storage snapshots. The physical standby database begins as a full copy of the primary database. When you are ready to create storage snapshots, sparse data files are created pointing to the full database files to apply redo shipped from the primary. These sparse files are then used in the standby database instance to apply redo. You can also open the sparse data files in Active Data Guard mode to supply read only access of current data.

When additional snapshots are required at different points in time, you repeat the process of creating new sparse files on top of the previously created sparse files to apply redo and keep the data current. This allows you to use a single full copy of the data files to use as multiple test masters from different points in time. Also, you can create a new test master in a matter of minutes because you do not have to drop the existing snapshots.

The following tasks assume that a physical standby database has already been created to be used as the source for the test masters. The database can be a container database (CDB) or non-container database (non-CDB).

- **Task 1: Prepare the Standby Database to Be Used as a Sparse Test Master**
- **Task 2: Configure the Sparse Test Master and Sparse Files on the Standby Site** In this task you convert the standby into a test master and create sparse files to receive and apply redo from the primary.
- **Task 3: Create Full Database Snapshots Using the New Sparse Test Master**
- **Task 4: Create a New Sparse Test Master Using a Previously Created Sparse Test Master** Create a new set of snapshots to provide a new test master and new sparse files to contain changes from the primary.
Related Topics

- **Using an Oracle Data Guard Standby Database as the Test Master**
  If the test master is a complete database that needs to be refreshed regularly, Oracle recommends creating the test master database as an Oracle Data Guard physical standby dedicated to this purpose.

- **Creating a PDB in a Primary Database**

**10.7.7.1 Task 1: Prepare the Standby Database to Be Used as a Sparse Test Master**

The existing files for the standby database are used to support snapshots. You create a series of sparse files pointing to the existing files of the standby. Redo received from the primary database is applied to these files. These sparse files allow the standby database to be used as a sparse test master and also kept current with the primary database.

1. **Stop redo apply at the standby.**
   To ensure that the structure of the database is at a quiesced state for creating the supporting files to build snapshots, redo apply should be turned off at the standby database.

   DGMGRL> edit database tm_standby set state='APPLY-OFF';

2. **Prepare the current standby database to be used as a test master.**
   The system must be configured as follows:

   a. **The disk group containing the database files must have the access_control.enabled attribute set to TRUE.**
      As SYSASM, log into Oracle ASM using SQL*Plus and configure the disk group.
      For example:

      SQL> alter diskgroup DATA set attribute 'ACCESS_CONTROL.ENABLED'='TRUE';

   b. **The operating system (OS) user of the database owner must be added as an explicit user of the disk group containing the database files.**
      For example:

      SQL> alter diskgroup DATA add user 'scott';

   c. **The database files that are to be used must have explicit permissions granted to the database owner OS user.**
      You must perform this step for all OS users that will be creating snapshots using these files and for all files that will be referenced by the snapshots.
      The following script can be used to build SQL statements to configure the ownership settings. Run the script while connected to the standby database.
using SQL*Plus. If the standby is a container database (CDB) you must be connected to the cdb$root container:

```sql
set newpage 0
set linesize 999
set pagesize 0
set feedback offset heading off
set echo off
set space 0
set tab off
set trimspool on
spool set_owner.sql
select 'ALTER DISKGROUP DATA set ownership
owner='.$'||'$'scott'||'$'||'
for file $'||'$'name'||'$'||';' from v$datafile;
exit
```

After running the previous script, log in to Oracle ASM using SQL*Plus as the SYSASM user, and run the commands in `set_owner.sql`.

```sql
SQL> @set_owner
```

3. Create a backup of the controlfile.

All snapshots are created using the current state of the standby database, so they need to know all of the files that make up the standby. Create a binary backup of the control file to allow future creation of the CREATE CONTROLFILE script required for additional snapshots.

```sql
SQL> ALTER DATABASE BACKUP CONTROLFILE TO '/home/oracle/snap_tm/control_tm.ctl';
```

4. Create the `rename_files.sql` script to create the sparse data files for the snapshot.

This script builds a series of RENAME statements to create the sparse data files to be used for the snapshot to apply redo received from the primary. Use a SQL statement similar to the following. Note that this statement uses the same directory structure as the original files, but the files will be created in the SPARSE disk group. The new file names will be created replacing '.' (periods) with '_' (underscores).

```sql
set newpage 0
set linesize 999
set pagesize 0
set feedback off
set heading off
set echo offset space 0
set tab off
set trimspool on
spool rename_files.sql
select 'EXECUTE dbms_dnfs.clonedb_renamefile ('||''||'name'||''||','||'replace(replace(name,'.','_'),''DATA/'','SPARSE/')'||'');'
```
This script produces output similar to the following:

```sql
EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TM_STANDBY/DATAFILE/
75939', '+SPARSE/TM_STANDBY/DATAFILE/system_515_930475939');

EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TM_STANDBY/
429CE0836E0166ACE05382C8
E50A1154/DATAFILE/system_567_930475945', '+SPARSE/TM_STANDBY/
429CE0836E0166ACE05382C8
82C850A1154/DATAFILE/system_567_930475945');

EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TM_STANDBY/DATAFILE/
sysoaux.571.9304
75939', '+SPARSE/TM_STANDBY/DATAFILE/sysoaux_571_930475939');

EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TM_STANDBY/
429CE0836E0166ACE05382C8
E50A1154/DATAFILE/sysoaux.516.930475945', '+SPARSE/TM_STANDBY/
429CE0836E0166ACE05382C8
82C850A1154/DATAFILE/sysoaux_516_930475945');

EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TM_STANDBY/DATAFILE/
undotbs1.497.930475939', '+SPARSE/TM_STANDBY/DATAFILE/undotbs1_497_930475939');

EXECUTE dbms_dnfs.clonedb_renamefile ('+DATA/TM_STANDBY/
429CE0836E0166ACE05382C8
E50A1154/DATAFILE/undotbs1.564.930475945', '+SPARSE/TM_STANDBY/
429CE0836E0166ACE05382C8
5382C850A1154/DATAFILE/undotbs1_564_930475945');
```

5. **Using ASMCMD, create directories for all of the directories identified in the script `rename_files.sql`.

When the `dbms_dnfs.clonedb_renamefile` function is executed, it requires that all directory structures used for the files already exist in ASM.

Use the output from the previous step to determine the structures required and then create them as needed. You can use ASMCMD to create the directories as in the following example:

```
cd ASMCMD [+> cd sparse
ASMCMD [+sparse] > ls
ASMCMD [+sparse] > mkdir tm_standby
ASMCMD [+sparse] > cd tm_standby
ASMCMD [+sparse/tm_standby] > mkdir datafile
ASMCMD [+sparse/tm_standby] > mkdir 429DC0E1BCBD1B90E05382C8E50A8E80
ASMCMD [+sparse/tm_standby] > mkdir 429CE0836E0166ACE05382C8E50A1154
ASMCMD [+sparse/tm_standby] > cd 429DC0E1BCBD1B90E05382C8E50A8E80
ASMCMD [+sparse/tm_standby/429DC0E1BCBD1B90E05382C8E50A8E80] > mkdir datafile
```
### 10.7.7.2 Task 2: Configure the Sparse Test Master and Sparse Files on the Standby Site

In this task you convert the standby into a test master and create sparse files to receive and apply redo from the primary database.

#### Note:

During this process you will not be creating a full snapshot database, you will be using portions of the existing standby database and adding sparse data files to the standby. The standby database controlfile will be modified to use the sparse files that are added. Going forward, the same standby instance is used, but redo apply will use the sparse files to store changes, leaving the original standby data files to serve as a sparse test master files for snapshots.

The existing data files are used to support full database snapshots with data as of the point in time the process was run.

1. Shutdown all instances of the TM_STANDBY database.

   `$ srvctl stop db -d tm_standby -o abort`

2. Using SQL*Plus, start one of the TM_STANDBY instances in mount mode.

   `SQL> startup mount`

3. Change the `DB_CREATE_FILE_DEST` setting in the standby instance to point to the SPARSE disk group.

   This ensures all new data files that are created will reside in the SPARSE disk group. To perform this step you must disable standby file management.

   `SQL> alter system set standby_file_management='MANUAL';`  
   `SQL> alter system set db_create_file_dest='+SPARSE';`

4. Run the `rename_files.sql` script created in Task 1 against the standby database.

   Running the script renames the data files in the TM_STANDBY controlfile and creates the sparse files for the snapshot.

   `SQL> @rename_files`

5. Re-enable standby file management.
Completing this step ensures that all new data files added to the primary will automatically be created by the standby when it receives the redo to create the data file.

```sql
SQL> alter system set standby_file_management='AUTO';
```

6. **Enable redo apply on TM_STANDBY.**
   Completing this step applies redo to the snapshot, keeping it current and preparing for the next round of snapshot creation.

   ```sql
   DGMGRL> edit database tm_standby set state='APPLY-ON';
   ```

7. **Restart the remaining instances in mount mode.**

   ```bash
   $ srvctl start db -d tm_standby -o mount
   ```

   **Note:**
   If you plan to clone a local pluggable database (PDB) at the primary database, then enable Active Data Guard mode at the standby. This requires an Active Data Guard license.

---

**Figure 10-15** Configuration with Test Master Files and Sparse Files for Redo Apply

![Configuration Diagram](image)

---

10.7.7.3 Task 3: Create Full Database Snapshots Using the New Sparse Test Master

At this time you can create full snapshots against the original files of the standby database as described in Creating a Snapshot of a Full Database.

You must use the backup controlfile created in step 2 of Task 1: Prepare the Standby Database to Be Used as a Sparse Test Master to build the CREATE CONTROLFILE statement. To use the file you can create a temporary database instance to mount the controlfile and execute the backup controlfile to trace command.

1. **Create a small PFILE file for the instance to use.**
At a minimum, the PFILE file should contain the following parameters:

```sql
control_files='/home/oracle/snap_tm/control_tm.ctl'  # This should be the control file name created above
db_name=primary  # This should be the db_name used in the Data Guard configuration
db_unique_name=temp  # This should be a unique name for a database instance on this host
sga_target=5g  # Provide enough memory to start the instance
```

2. Set your environment to point to a unique ORACLE_SID.

```bash
$ export ORACLE_SID=temp
```

3. Using SQL*Plus, start the instance in mount mode using the PFILE created in Step 1.

```sql
SQL> startup mount pfile='/home/oracle/snap_tm/pfile.ora'
```

4. Build the create controlfile statement and the rename files script.

   Use steps 1 and 2 in Creating a Snapshot of a Full Database to build the CREATE CONTROLFILE statement and the rename files script. The rename files script created in step 4 of Task 1: Prepare the Standby Database to Be Used as a Sparse Test Master can be used, but you must modify the directory structure of the sparse files to be created.

![Figure 10-16 Configuration After Creating Snapshots](image)

10.7.7.4 Task 4: Create a New Sparse Test Master Using a Previously Created Sparse Test Master

Create a new set of snapshots to provide a new test master and new sparse files to contain changes from the primary.

It is possible that at periodic intervals you will want to use the standby database to create additional snapshots without having to build a complete copy of the test master.
You can repeat the process performed in the previous three tasks to do just that, taking advantage of the hierarchical snapshot functionality. The new test master is built on top of the latest existing snapshot that is applying redo. This snapshot becomes read-only and a new snapshot is built to continue the redo apply processing.

Do the following to configure the standby for the new test master time line:

1. Perform the steps from **Task 1: Prepare the Standby Database to Be Used as a Sparse Test Master** with the following changes:

   The procedure is generally the same, however in this case the test master already uses sparse files.

   a. In Step 2, change all the commands that alter the DATA disk group to instead alter the SPARSE disk group.

      If the SPARSE disk group is already appropriately configured, you can skip the corresponding `ALTER DISKGROUP ... SET ATTRIBUTE` or `ALTER DISKGROUP ... ADD USER` command.

   b. In Step 4, supply a different name for the snapshot.

      You are creating a new snapshot, so the files need unique names from what was previously used. As a suggestion, you can append an identifier to the end of the file name to help identify it with the snapshot to be built. For example, if this was the original command:

      ```sql
      EXECUTE dbms_dnfs.clonedb_renamefile ('+SPARSE/TM_STANDBY/DATAFILE/system_515_930475939', '+SPARSE/TM_STANDBY/DATAFILE/system_515_930475939');
      ```

      You can add an identifier to the end of the file name to create a unique file name, as shown here:

      ```sql
      EXECUTE dbms_dnfs.clonedb_renamefile ('+SPARSE/TM_STANDBY/ DATAFILE/system_515_930475939', '+SPARSE/TM_STANDBY/DATAFILE/ system_515_930475939_Dec_15_16');
      ```

      This step must be repeated for each statement for the `rename_files.sql` script.

2. Perform the steps as described in **Task 2: Configure the Sparse Test Master and Sparse Files on the Standby Site**.
The process to create a new sparse test master can be repeated up to 9 times, which would create an environment 10 levels deep with the original standby data files and 9 hierarchical snapshots. When repeating the process for the ninth time, do not create a new snapshot to receive the redo from the primary database.

If you reach the maximum 10 levels, you have multiple options:
If you have enough space to maintain multiple copies of the standby database and snapshots, then create a new tree of hierarchical snapshots based on another standby database. The original standby data files and snapshots can remain as long as required.

If you do not have enough space to maintain multiple copies of the standby database and snapshots, then:

1. Delete the snapshots and their associated data files, including any snapshots used as test masters.
2. Refresh the standby.
3. Create a new tree of hierarchical snapshots.

Create a new standby database on a different environment and create a new tree of hierarchical snapshots.

Related Topics

- Task 1: Prepare the Standby Database to Be Used as a Sparse Test Master
- Task 2: Configure the Sparse Test Master and Sparse Files on the Standby Site

In this task you convert the standby into a test master and create sparse files to receive and apply redo from the primary database.

10.7.8 Creating Sparse Test Masters for PDBs

This procedure creates a hierarchical snapshot tree or sparse test masters manually for a pluggable database (PDB) in an Oracle Multitenant database.

The test master must be closed while making the daily reference snapshot. The downtime is very short (less than 5 minutes). You can use a replication mechanism, such as Oracle GoldenGate, to keep the sparse test master current with the production PDB. For more information about configuring Oracle GoldenGate with PDBs, see Configuring Oracle GoldenGate in a Multitenant Container Database in Oracle GoldenGate Oracle Installation and Setup Guide. The following example assumes you are using Oracle GoldenGate.

STEP 1: Create First Test Master PDB From the PROD PDB

This is a traditional PDB clone operation to instantiate the test master PDB. Once the clone completes, you configure Oracle GoldenGate to extract changes from the PRODPDB1 PDB in production and replicate these changes to the test master TMPDB1 PDB.

1. Run the following commands on the PROD container database (CDB) root:

```
PRODCDB> alter pluggable database prodpdb1 close;
PRODCDB> alter pluggable database prodpdb1 open read only;
```

2. Run the following commands from the test master CDB root:

```
TMCDB> create database link PROD_DBLINK
       connect to system identified by password using 'PROD_CDB';

TMCDB> create pluggable database TMPDB1
       from PRODPDB1@PROD_DBLINK;
```
3. Configure Oracle GoldenGate so that the changes made at the PRODPDB1 PDB will be extracted, replicated and applied to the TMPDB1 PDB. After configuring the extract and replicate and starting the extract process, open PRODPDB1 PDB in read write mode.

**Note:**
The PRODPDB1 PDB cannot be opened for changes until after Oracle GoldenGate has been configured and the extract process started.

PRODCDB> alter pluggable database PRODPDB1 close;
PRODCDB> alter pluggable database PRODPDB1 open;

At this point, you have a full copy of PRODPDB1 on test master as TMPDB1 receiving all data changes made at PRODPDB1.

**Note:**
Oracle GoldenGate does not replicate data dictionary changes such as CREATE TABLESPACE or ADD DATAFILE. Only schema changes are replicated from PRODPDB1 to TMPDB1.

Although TMPDB1 can be opened in read/write mode, you should leave it in read-only mode because the only changes it should receive are from PRODPDB1 through Oracle GoldenGate.
To create a snapshot from TMPDB1, the test master PDB must be opened in read-only mode. To provide a test master PDB from which you can create snapshots, and a test master PDB that is kept current with its source, you need two PDBs. The next step shows how you can accomplish this.

**STEP 2: Create Daily Read-Only Snapshot and Move TMPDB1 PDB to a New Sparse Test Master PDB**

This step creates a read-only snapshot PDB that acts as a test master. You can then create read/write snapshot PDBs from this read-only snapshot PDB every day. The main steps are:

- Create a (daily) read-only snapshot PDB that you can make available to private read/write clients.
- Create a new sparse TMPDB1 PDB pointing back to the read-only daily snapshot PDB. The new TMPDB1 PDB also accepts and applies changes from PRODPDB1.

Connect to the TMPDB1 PDB, then run the following commands:

```
TMCDB> alter session set container = CDB$ROOT;

# Stop the Oracle GoldenGate replicat process at the Test Master database. This allows all changes made at PRODPDB1 to continue to be extracted and then applied to TMPDB1 when the replicat process is restarted.

# Close the test master PDB.
TMCDB> alter pluggable database TMPDB1 close;

# Write the test master PDB metadata to an XML file.
TMCDB> alter pluggable database TMPDB1 unplug into '/home/oracle/snapshot/TMPDB1_monday.XML';

# Drop the test master PDB, but keep the data files.
TMCDB> drop pluggable database TMPDB1 keep datafiles;

# Create a TMPDB1_MONDAY PDB using the XML file you just created. Use the NOCOPY clause to reuse the original data files.
TMCDB> create pluggable database TMPDB1_MONDAY using '/home/oracle/snapshot/TMPDB1_monday.XML' nocopy;

# Open the new TMPDB1_MONDAY PDB. The PDB must be opened once in read/write mode to complete the creation process.
TMCDB> alter pluggable database TMPDB1_MONDAY open;
TMCDB> alter pluggable database TMPDB1_MONDAY close;
TMCDB> alter pluggable database TMPDB1_MONDAY open read only;

# Create the new TMPDB1 PDB to receive changes from PRODPDB1. This PDB must have the same name as the original test master PDB to ensure no changes are required to the Oracle GoldenGate configuration.
TMCDB> create pluggable database TMPDB1 from TMPDB1_MONDAY
```
create_file_dest=’+SPARSE’

snapshot copy;

# Open the new TMPDB1 PDB. The PDB must be opened once in read/write
# mode to complete the PDB creation process.

TMCDB> alter pluggable database TMPDB1 open;
TMCDB> alter pluggable database TMPDB1 close;
TMCDB> alter pluggable database TMPDB1 open read only;

# Restart the Oracle GoldenGate replicat process to the new TMPDB1
# PDB. The Oracle GoldenGate replicat process now applies changes from
# PRODPDB1 to the TMPDB1 snapshot and all changes are written to
# sparse files.

The following figure shows the TMPDB1 created from TMPDB1_MONDAY. The original
TMPDB1 has been renamed to TMPDB1_Monday as part of the DROP PLUGGABLE
DATABASE/CREATE PLUGGABLE DATABASE steps listed above. The new TMPDB1 is a
sparse snapshot pluggable database that, until any changes are made to TMPDB1,
looks exactly like TMPDB1_Monday. Oracle GoldenGate applies redo to the new
TMPDB1 snapshot without having to make any changes to the replicat configuration

**Figure 10-20  TMPDB1 Created From TMPDB1_MONDAY**

STEP 3: Create Read/Write Snapshot From TMPDB1_MONDAY

You create the snapshots from TMPDB1_MONDAY, not from TMPDB1. This allows TMPDB1 to
continue receiving and applying changes from PRODPDB1.
Connect to the TMPDB1_MONDAY PDB, then run the following commands:

```
TMCDB> alter session set container = cdb$ROOT;
TMCDB> create pluggable database TEST_MONDAY_JIM from TMPDB1_MONDAY
        create_file_dest='+SPARSE'
        snapshot copy;
TMCDB> alter pluggable database TEST_MONDAY_JIM open;
```

The following figure shows the TEST_MONDAY_JIM snapshot PDB created from TMPDB1_MONDAY. TEST_MONDAY_JIM uses TMPDB1_MONDAY as its parent so all data in TMPDB1_MONDAY_JIM is that same as the data in TMPDB1_MONDAY until Jim begins making changes to his snapshot PDB. Oracle GoldenGate continues to receive and apply redo to TMPDB1.

**Figure 10-21**   TEST_MONDAY_JIM Created From TMPDB1_MONDAY

When you need to create another test master and snapshot, you just need to repeat Step 2. For example, to create a test master on Tuesday, you can do the following:

Start a SQL*Plus session for the TMPDB1 PDB.

```
TMCDB> alter session set container = CDB$ROOT;
```
# Stop the Oracle GoldenGate replicat process from applying changes to
# TMPDB1

TMCD> alter pluggable database TMPDB1 close;

# Close the test master PDB
TMCD> alter pluggable database TMPDB1 close;

# Write the test master PDB metadata to an XML file
TMCD> alter pluggable database TMPDB1 unplug into
'/home/oracle/snapshots/TMPDB1_tuesday.XML';

# Drop the test master PDB, but keep the data files
TMCD> drop pluggable database TMPDB1 keep datafiles;

# Create a TMPDB1_TUESDAY PDB from the XML file
TMCD> create pluggable database TMPDB1_TUESDAY using
'/home/oracle/snapshots/TMPDB1_tuesday.XML' nocopy;

# Open the new TMPDB1_TUESDAY PDB
TMCD> alter pluggable database TMPDB1_TUESDAY open;
TMCD> alter pluggable database TMPDB1_TUESDAY close;
TMCD> alter pluggable database TMPDB1_TUESDAY open read only;

# Create the new TMPDB1 PDB as a snapshot PDB
TMCD> create pluggable database TMPDB1 from TMPDB1_TUESDAY
    create_file_dest='+SPARSE'
    snapshot copy;

# Open the TMPDB1 PDB
TMCD> alter pluggable database TMPDB1 open;
TMCD> alter pluggable database TMPDB1 close;
TMCD> alter pluggable database TMPDB1 open read only;

# Restart the Oracle GoldenGate replicat process to apply changes to
# the new TMPDB1

You can now create read/write snapshot PDBs from TMPDB1_TUESDAY, similar to
Step 3 above. The same as with full database sparse test masters, you can repeat
this process up to 9 times in total before needing to either create a new TMPDB1 test
master or drop and recreate the original TMPDB1 to begin building a new hierarchical
snapshot tree.

10.7.9 Determining All Snapshots Associated with a Test Master

Use this query to discover all of the children associated with a test master.

Consider the following configuration for a test master with multiple children.
You can use a query of the associated SYSTEM data files for each database to list all the children within the same tree. The query selects just the SYSTEM data file for each database or PDB. The data file must exist for all clones and parents, and there should be only 1 data file for each. The `START WITH` clause provides a starting point of a file that is not a cloned file, which is the original test master parent.

Connect to the Oracle ASM instance and run this command as the SYSASM user.

```sql
SELECT clonefilename "Child", snapshotfilename "Parent"
FROM v$clonedfile
WHERE LOWER(snapshotfilename) LIKE '%system.%'
START WITH snapshotfilename NOT IN (SELECT clonefilename FROM v$clonedfile)
CONNECT BY LOWER(clonefilename) = PRIOR (snapshotfilename);
```

The results of this query for database-based snapshots would be similar to the following:

Child
----
Parent
-----------------------------
+SPARSE/SPAP001/DATAFILE/SYSTEM.256.1011532891
 +DATA/TESTMASTER/DATAFILE/system.270.1011530981
+SPARSE/SPAP002/DATAFILE/SYSTEM.265.1011532969
 +DATA/TESTMASTER/DATAFILE/system.270.1011530981
+SPARSE/SPAP1011/DATAFILE/SYSTEM.270.1011533005
 +SPARSE/SPAP001/DATAFILE/system.256.1011532891
+SPARSE/SPAP1012/DATAFILE/SYSTEM.275.1011780925

---
If you created folders in Oracle ASM that contained the database name, as shown in the above result, then the database name in the `CLONEFILENAME` string is the snapshot, and the database name in the `SNAPSHOTFILENAME` string is the master for that snapshot.

The results of this query for PDB-based snapshots would be similar to the following:

<table>
<thead>
<tr>
<th>CLONEFILENAME</th>
<th>SNAPSHOTFILENAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>+SPARSEC1/CDB001/8BDBC355D43721F5E053412E850AB5D1/DATAFILE/SYSTEM.256.1011532891</td>
<td>+DATAC1/CDB001/8BDBC355D42D21F5E053412E850AB5D1/DATAFILE/system.270.1011530981</td>
</tr>
<tr>
<td>+SPARSEC1/CDB001/8BDBC355D43E21F5E053412E850AB5D1/DATAFILE/SYSTEM.265.1011532969</td>
<td>+DATAC1/CDB001/8BDBC355D44821F5E053412E850AB5D1/DATAFILE/system.270.1011533005</td>
</tr>
<tr>
<td>+SPARSEC1/CDB001/8BDBC355D44D21F5E053412E850AB5D1/DATAFILE/SYSTEM.281.1011781103</td>
<td></td>
</tr>
</tbody>
</table>

In this case, the folder name in Oracle ASM is the GUID associated with the PDB. To determine the name of each snapshot PDB and its master, you must do the following:

1. Log in to the CDB that has the name shown in the results, for example, `CDB001`.
2. Run a query against the `CDB_PDBS` view to translate the GUIDs into the PDB names, as shown below:

```sql
SELECT pdb_name, guid FROM CDB_PDBS
WHERE guid IN
    ('8BDBC355D42D21F5E053412E850AB5D1', '8BDBC355D43721F5E053412E850AB5D1',
     '8BDBC355D44821F5E053412E850AB5D1', '8BDBC355D44D21F5E053412E850AB5D1',
     '8BDBC355D44021F5E053412E850AB5D1');
```
Then use this information to determine the parent/child relationship among the PDBs in the original query results.

### 10.7.10 Doing a Sparse Copy

The ASM `cp` command copies a sparse file to a new destination. However, this operation copies the sparse file with all the blocks instantiated from the parent. The “sparse copy” feature enables you to do a sparse copy of a file.

You can have multiple ASM instances running at the same time. If an operation involves a source or a destination on a different ASM instance other than the one it is executing on, it is treated as a remote ASM instance. You can do a sparse copy on a local ASM instance, or between a local and a remote ASM instance. However, sparse copy does not work between two remote ASM instances.

To do a sparse copy, you use the new `--sparse` option in the existing ASM `cp` command. The syntax looks like the following:

```bash
ASMCMD> cp --sparse <src_sparse_file> <tgt_file>
```

A new ASM command called `setsparseparent` enables you to set the parent of a sparse file. If you do a sparse copy of a file to a sparse destination on a local ASM instance, its parent is set as part of the sparse copy operation. However, if the destination is on a remote ASM instance, you have to set its parent explicitly using the `setsparseparent` command.

The `setsparseparent` command requires sparse child file and parent file as parameters. It sets the parent of the sparse child file to the new parent file. The syntax looks like the following:

```bash
ASMCMD> setsparseparent <sparse_file> <parent_file>
```

The `cp` ASM command performs the following validations before doing a sparse copy operation. The operation is allowed only if it satisfies the following criteria:

- The source file must exist and must be a sparse file.
- If you specify multiple source sparse files, all of them must be on the same ASM instance.
- Copying multiple sparse files on a remote ASM instance to a destination on a local ASM instance and vice versa is allowed provided all source files are on the same ASM instance.
- Destination file should be backed by a sparse disk group. However, it can be a non-sparse file if event “KFTST_KFPKG_CP_SPARSE” is set. This event is...
required to validate sparse copy operation by merging and copying the files to a non-sparse destination.

- Both source and destination cannot be on a remote ASM instance. However, either source or destination can be on a remote ASM instance.
- If the destination is on a remote ASM instance, its file type cannot be validated and you have to ensure that it is backed by a sparse disk group. You also have to set the parent explicitly using the ASM `setsparseparent` command.
- If the destination is a non-sparse file and you run the `setsparseparent` command, the command will fail because the child file should be sparse. This is a second-level validation if the destination is a non-sparse file.

The `setsparseparent` ASM command performs the following validations before it sets the parent. The operation is allowed only if it satisfies the following criteria:

- The child file must exist and must be a sparse file.
- The parent file must exist. It can be a sparse or a non-sparse file.
- Parent and child files must be present on same ASM instance.

![Note:]

You have to ensure that the files you specify in the `setsparseparent` ASM command have a valid parent-child relationship. The command cannot perform this check for files on remote ASM instances. If the files do not have a valid parent-child relationship, then data integrity and corruption issues are likely to occur.

Example 1: The following ASM command copies sparse file “TBS_1.264.908376549” to the destination “+SPARSEDG/child_1”.

```
ASMCMD> cp --sparse +SPARSEDG/MERGE/DATAFILE/TBS_1.264.908376549 +SPARSEDG/child_1
```

Example 2: The following ASM command sets parent “tbs_1.269.908374993” for the sparse file “remote_child_10”.

```
ASMCMD> setsparseparent +SPARSEDG/remote_child_10 +DATAFILE/DATAFILE/tbs_1.269.908374993
```

Example 3: The following command copies sparse child files child_1, child_2 and child_3 to the destination directory +SPARSEDG.

```
ASMCMD> cp --sparse +SPARSEDG/DATAFILE/child_1 +SPARSEDG/DATAFILE/child_2 +SPARSEDG/DATAFILE/child_3 +SPARSEDG/
```

### 10.8 Managing Sparse Griddisks

You can resize, recreate, or monitor the activity of sparse griddisks.

- **Resizing the Virtual Space**

  When $VASM_DISKGROUP.FREE_MB or $VASM_DISK.FREE_MB is running low, you need to increase the virtual address space.
• **Resizing the Physical Space**
  When the grid disks are running out of physical space, then you need to increase the physical size of the grid disks.

• **Monitoring Sparse Disk Group Utilization**

• **Repurposing Sparse Griddisks**
  You can change sparse griddisks back to normal griddisks.

### 10.8.1 Resizing the Virtual Space

When $V$\$ASM\_DISKGROUP.FREE\_MB or $V$\$ASM\_DISK.FREE\_MB is running low, you need to increase the virtual address space.

1. To increase the size of the virtual space:
   a. Run the following command on the cells, specifying all the grid disks for the SPARSE disk group:

   ```
   CellCLI> alter griddisk SPARSE\_CD\_00\_CELL01,SPARSE\_CD\_01\_CELL01,.....,SPARSE\_CD\_11\_CELL01 virtualSize=newSize
   ```

   For example, on the first cell:

   ```
   CellCLI> alter griddisk SPARSE\_CD\_00\_CELL01,SPARSE\_CD\_01\_CELL01,SPARSE\_CD\_02\_CELL01,SPARSE\_CD\_03\_CELL01,SPARSE\_CD\_04\_CELL01,SPARSE\_CD\_05\_CELL01,SPARSE\_CD\_06\_CELL01,SPARSE\_CD\_07\_CELL01,SPARSE\_CD\_08\_CELL01,SPARSE\_CD\_09\_CELL01,SPARSE\_CD\_10\_CELL01,SPARSE\_CD\_11\_CELL01 virtualSize=12000G
   ```

   GridDisk SPARSE\_CD\_00\_CELL01 successfully altered
   GridDisk SPARSE\_CD\_01\_CELL01 successfully altered
   ...

   For example, on the next cell:

   ```
   CellCLI> alter griddisk SPARSE\_CD\_00\_CELL02,SPARSE\_CD\_01\_CELL02,SPARSE\_CD\_02\_CELL02,SPARSE\_CD\_03\_CELL02,SPARSE\_CD\_04\_CELL02,SPARSE\_CD\_05\_CELL02,SPARSE\_CD\_06\_CELL02,SPARSE\_CD\_07\_CELL02,SPARSE\_CD\_08\_CELL02,SPARSE\_CD\_09\_CELL02,SPARSE\_CD\_10\_CELL02,SPARSE\_CD\_11\_CELL02 virtualSize=12000G
   ```

   GridDisk SPARSE\_CD\_00\_CELL02 successfully altered
   GridDisk SPARSE\_CD\_01\_CELL02 successfully altered
   ...
Note:
You must change the size of the grid disks on all cells before making any changes in Oracle ASM.

b. On an ASM instance, resize the disk group to this new size:

```
SQL> alter diskgroup SPARSE resize all size newSize;
```

For example:

```
SQL> alter diskgroup SPARSE resize all size 12000G;
```

2. To decrease the size of the virtual space:

a. On an ASM instance, resize the disk group to this new size:

```
SQL> alter diskgroup SPARSE resize all size newSize;
```

For example:

```
SQL> alter diskgroup SPARSE resize all size 8000G;
```

b. Run the following command on the cells, specifying all the grid disks for the SPARSE disk group:

```
CellCLI> alter griddisk SPARSE_CD_00_CELL01,SPARSE_CD_01_CELL01,...,SPARSE_CD_11_CELL01 virtualSize=newSize
```

For example, on the first cell:

```
CellCLI> alter griddisk  
SPARSE_CD_00_CELL01,SPARSE_CD_01_CELL01,SPARSE_CD_02_CELL01,SPARSE_CD_03_CELL01,SPARSE_CD_04_CELL01,SPARSE_CD_05_CELL01,SPARSE_CD_06_CELL01,SPARSE_CD_07_CELL01,SPARSE_CD_08_CELL01,SPARSE_CD_09_CELL01,SPARSE_CD_10_CELL01,SPARSE_CD_11_CELL01 virtualSize=8000G
```

GridDisk SPARSE_CD_00_CELL01 successfully altered
GridDisk SPARSE_CD_01_CELL01 successfully altered ...

For example, on the next cell:

```
CellCLI> alter griddisk  
SPARSE_CD_00_CELL02,SPARSE_CD_01_CELL02,SPARSE_CD_02_CELL02,SPARSE_CD_03_CELL02,SPARSE_CD_04_CELL02,SPARSE_CD_05_CELL02,SPARSE_CD_06_CELL02,SPARSE_CD_07_CELL02,SPARSE_CD_08_CELL02,SPARSE_CD_09_CELL02
```

GridDisk SPARSE_CD_00_CELL02 successfully altered...
10.8.2 Resizing the Physical Space

When the grid disks are running out of physical space, then you need to increase the physical size of the grid disks.

You can determine the amount of physical space left by comparing the values of the \texttt{TOTAL\_MAT\_MB} and \texttt{ALLOCATED\_MAT\_MB} columns in \texttt{V\_ASM\_DISK\_SPARSE}. When the values in these two columns are close in size, then you need to increase the physical size of the grid disks.

1. To increase the size of the physical disk space:
   
a. Before increasing the physical size of grid disks ensure there is free space available on the respective celldisks.

   ```
   [root@exa01adm01 tmp]# dcli -g ~/cell_group -l root "cellcli -e list celldisk \n   attributes name,freespace"
   exa01celadm01: CD_00_exa01celadm01 0
   exa01celadm01: CD_01_exa01celadm01 0
   exa01celadm01: CD_02_exa01celadm01 0
   exa01celadm01: CD_03_exa01celadm01 0
   exa01celadm01: CD_04_exa01celadm01 0
   exa01celadm01: CD_05_exa01celadm01 0
   exa01celadm01: CD_06_exa01celadm01 0
   exa01celadm01: CD_07_exa01celadm01 0
   exa01celadm01: CD_08_exa01celadm01 0
   exa01celadm01: CD_09_exa01celadm01 0
   exa01celadm01: CD_10_exa01celadm01 0
   exa01celadm01: CD_11_exa01celadm01 0
   ...
   ```

   If there is no available free space, then you have to free up disk space being used by the Oracle ASM disk groups.

   b. Run the following command on the cells, specifying the grid disk to resize:

   ```
   CellCLI> alter griddisk
data01\_CD\_00\_exa01celadm01,data01\_CD\_01\_exa01celadm01,...
data01\_CD\_11\_exa01celadm01 size=newPhysicalSize
   ```
For example, on the first cell:

```
CellCLI> alter griddisk
data01_CD_00_exa01celadm01,data01_CD_01_exa01celadm01,
data01_CD_02_exa01celadm01,data01_CD_03_exa01celadm01,data01_CD_04_exa01celadm01,
data01_CD_05_exa01celadm01,data01_CD_06_exa01celadm01,data01_CD_07_exa01celadm01,
data01_CD_08_exa01celadm01,data01_CD_09_exa01celadm01,data01_CD_10_exa01celadm01,
data01_CD_11_exa01celadm01 size=12000G
```

And then on the next cell:

```
CellCLI> alter griddisk
data01_CD_00_exa01celadm02,data01_CD_01_exa01celadm02,
data01_CD_02_exa01celadm02,data01_CD_03_exa01celadm02,data01_CD_04_exa01celadm02,
data01_CD_05_exa01celadm02,data01_CD_06_exa01celadm02,data01_CD_07_exa01celadm02,
data01_CD_08_exa01celadm02,data01_CD_09_exa01celadm02,data01_CD_10_exa01celadm02,
data01_CD_11_exa01celadm02 size=12000G
```

c. You do not need to do anything on the Oracle ASM instance.

2. To decrease the size of the physical disk space:

   a. You do not need to do anything on the Oracle ASM instance.

   b. Make sure you are not using more space in the Oracle ASM disk group than the size to which you are shrinking the physical disk space.

```
SQL> SELECT sum(allocated_mat_mb) FROM v$asm_disk_sparse
       WHERE group_number = group_number_of_diskgrp_to_shrink;
```

If the amount of used physical space is more than the size you plan to shrink the disk down to, then you must drop objects from the sparse disk group until the used space is below the new physical space threshold.

c. Shrink the size of the disks by running the following command on the cells, specifying the grid disks to resize:

```
CellCLI> alter griddisk
data01_CD_00_exa01celadm01,data01_CD_01_exa01celadm01,...
        ,data01_CD_11_exa01celadm01 size=newPhysicalSize
```

For example, on the first cell:

```
CellCLI> alter griddisk
data01_CD_00_exa01celadm01,data01_CD_01_exa01celadm01,
data01_CD_02_exa01celadm01,data01_CD_03_exa01celadm01,data01_CD_04_exa01celadm01,
data01_CD_05_exa01celadm01,data01_CD_06_exa01celadm01,data01_CD_07_exa01celadm01,
data01_CD_08_exa01celadm01,data01_CD_09_exa01celadm01,data01_CD_10_exa01celadm01,
data01_CD_11_exa01celadm01 size=12000G
```
data01_CD_08_exa01celadm01,data01_CD_09_exa01celadm01,data01_CD_10_exa01celadm01,
data01_CD_11_exa01celadm01 size=4000G

And then on the next cell:

CellCLI> alter griddisk
data01_CD_00_exa01celadm02,data01_CD_01_exa01celadm02,
data01_CD_02_exa01celadm02,data01_CD_03_exa01celadm02,data01_CD_04_exa01celadm02,
data01_CD_05_exa01celadm02,data01_CD_06_exa01celadm02,data01_CD_07_exa01celadm02,
data01_CD_08_exa01celadm02,data01_CD_09_exa01celadm02,data01_CD_10_exa01celadm02,
data01_CD_11_exa01celadm02 size=4000G

Related Topics
• Resizing Grid Disks
  You can resize grid disks and Oracle ASM disk groups to shrink one with excess free space and increase the size of another that is near capacity.

10.8.3 Monitoring Sparse Disk Group Utilization


The following table describes the columns in V$ASM_DISK_SPARSE:

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP_NUMBER</td>
<td>The number of the disk group containing the disk.</td>
</tr>
<tr>
<td>DISK_NUMBER</td>
<td>The number assigned to the disk within this disk group.</td>
</tr>
<tr>
<td>INCARNATION</td>
<td>The incarnation number for the disk.</td>
</tr>
<tr>
<td>ALLOCATED_MAT_MB</td>
<td>The total used physical and materialized capacity on the disk.</td>
</tr>
<tr>
<td>TOTAL_MAT_MB</td>
<td>The total physical capacity on the disk.</td>
</tr>
<tr>
<td>SPARSE_READS</td>
<td>The total number of I/O read requests on non-materialized regions of the disk.</td>
</tr>
<tr>
<td>SPARSE_BYTES_READ</td>
<td>The total number of bytes read from non-materialized regions of the disk.</td>
</tr>
<tr>
<td>SPARSE_READ_TIME</td>
<td>The time taken by sparse read I/O operations.</td>
</tr>
</tbody>
</table>

The following table describes the columns in V$ASM_DISKGROUP_SPARSE:
Table 10-2  V$ASM_DISKGROUP_SPARSE Columns and Descriptions

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP_NUMBER</td>
<td>The cluster-wide number assigned to the disk group.</td>
</tr>
<tr>
<td>ALLOCATED_MAT_MB</td>
<td>The total used physical and materialized capacity of the disk group.</td>
</tr>
<tr>
<td>TOTAL_MAT_MB</td>
<td>The total physical capacity of the disk group.</td>
</tr>
</tbody>
</table>

The following example shows the used (allocated) space and the total space for disks in a specific disk group:

```
SQL> select DISK_NUMBER dsk_num, ALLOCATED_MAT_MB alloc, TOTAL_MAT_MB total
    from V$ASM_DISK_SPARSE
    where GROUP_NUMBER = 5;
```

<table>
<thead>
<tr>
<th>DSK_NUM</th>
<th>ALLOC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5536</td>
<td>57336</td>
</tr>
<tr>
<td>1</td>
<td>5424</td>
<td>57336</td>
</tr>
<tr>
<td>2</td>
<td>5532</td>
<td>57336</td>
</tr>
<tr>
<td>3</td>
<td>5424</td>
<td>57336</td>
</tr>
<tr>
<td>4</td>
<td>5424</td>
<td>57336</td>
</tr>
</tbody>
</table>

In the following example, sparse ASM grid disks were created with an actual size of 56 GB and a virtual size of 560 GB. When you query V$ASM_DISK's OS_MB and TOTAL_MB columns, you can see the virtual size of 573440 MB (573440 MB / 1024 = 560 GB).

```
SQL> select os_mb, total_mb from v$asm_disk where group_number=4;
```

<table>
<thead>
<tr>
<th>OS_MB</th>
<th>TOTAL_MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>573440</td>
<td>573440</td>
</tr>
<tr>
<td>573440</td>
<td>573440</td>
</tr>
<tr>
<td>573440</td>
<td>573440</td>
</tr>
</tbody>
</table>

Querying V$ASM_DISK_SPARSE for TOTAL_MB, you can see the actual size of the ASM grid disk available for use. Note that each ASM grid disk stores metadata information of approximately 2 MB per 16 GB of space allocated to the sparse ASM grid disk. For 56 GB allocated per grid disk in this example, 8 MB of space is reserved for sparse disk metadata (57336 MB + 8 MB = 57344 MB / 1024 = 56 GB).

```
SQL> select total_mb from v$asm_disk_sparse where group_number=4;
```

<table>
<thead>
<tr>
<th>TOTAL_MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>57336</td>
</tr>
<tr>
<td>57336</td>
</tr>
<tr>
<td>57336</td>
</tr>
</tbody>
</table>
10.8.4 Repurposing Sparse Griddisks

You can change sparse griddisks back to normal griddisks.

If you previously created sparse griddisks, but now want to use them as normal griddisks, you can drop and recreate the disks.

1. Drop the snapshot database that is currently using the sparse griddisks.

   RMAN> startup mount force;
   RMAN> delete database;

2. Using SQL*Plus or ASMCMD, drop the Oracle ASM disk group that contains the sparse griddisks.

   SQL> DROP DISKGROUP sparse INCLUDING CONTENTS force;

3. Using CellCLI, drop the griddisks in the storage cells.

   cellcli -e drop griddisk all harddisk prefix=SPARSEC1

4. Recreate the griddisks.

   When you create the griddisks, use a similar size as the other disks and add back to the disk group of your choice. Do not specify the sparse attribute. See CREATE GRIDDISK for the command syntax.

5. Add the recreated griddisks to an Oracle ASM disk group.

   Use the SQL ALTER DISKGROUP command with the ADD DISK clause to add the disk to the Oracle ASM disk group using syntax similar to the following:

   SQL> ALTER DISKGROUP disk_group_name ADD DISK 'o/cell_IPaddress/data*';

Related Topics

- Creating a Sparse Grid Disk

10.9 Monitoring Exadata Snapshots Using Database Statistics and Wait Events

The following table describes database specific statistics that are useful for monitoring Exadata Snapshots. The statistics are available in various dynamic performance views, including V$SYSSTAT, and may be displayed in the Global Activity Statistics or Instance Activity Statistics section of an AWR report.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>physical read snap IO requests base</td>
<td>The number of physical I/Os on the parent or base file.</td>
</tr>
<tr>
<td>physical read snap IO requests copy</td>
<td>The number of physical I/Os on the snapshot file.</td>
</tr>
</tbody>
</table>
The following table describes specific database wait events that are useful for monitoring Exadata Snapshots. The wait events are visible in various dynamic performance views, including V$SESSION, V$SYSTEM_EVENT and V$SESSION_EVENT, and may be displayed in the Wait Event sections of the AWR report.

<table>
<thead>
<tr>
<th>Wait Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell physical read no I/O</td>
<td>The wait event appears when a read is done from a sparse disk group, and did not return any data.</td>
</tr>
</tbody>
</table>
Upgrading Oracle Exadata System Software

Oracle Exadata System Software, operating systems, and component firmware are upgraded and patched using Oracle utilities.

Patches include all required software updates, operating system updates, and firmware updates. Use the following utilities to upgrade and patch Oracle Exadata Database Machine:

- The `patchmgr` utility is used for Oracle Exadata System Software, and RDMA Network Fabric switches.
- The DB Node Update utility (`dbnodeupdate.sh` utility) is used for database servers.

In addition, Oracle recommends running the Oracle EXAchk utility before and after performing planned maintenance.

Upgrade and patching information is available from My Oracle Support note 888828.1. Additional information is available in the patch README, and associated support notes.

For detailed information about updating the Oracle Exadata System Software, operating systems, and component firmware, refer to Updating Exadata Software.

Related Topics

- Exadata Database Machine and Exadata Storage Server Supported Versions (My Oracle Support Doc ID 888828.1)
Installation Information

This appendix contains information about Oracle Exadata Storage Server installation. Refer to this installation information if the Oracle Exadata Storage Server hardware and software are not already installed.

- **Installation of Oracle Exadata Storage Server**
  This section describes the installation process for a storage server.

- **Background Processes in the Oracle Exadata Storage Server Environment**

Related Topics

- **Configuring Oracle Exadata System Software**
  This chapter describes the major steps to configure Oracle Exadata System Software.

## B.1 Installation of Oracle Exadata Storage Server

This section describes the installation process for a storage server.

- **Installing Oracle Exadata Storage Server Hardware**
  Use these steps to complete the hardware installation necessary for standalone Oracle Exadata Storage Servers connecting to Oracle Exadata Database Machine.

- **Installing the Oracle Exadata System Software on a Cell**
  This procedure describes how to install the Oracle Exadata System Software on an Oracle Exadata Storage Server.

- **Updates Made During Installation**

- **Online Documentation**

## B.1.1 Installing Oracle Exadata Storage Server Hardware

Use these steps to complete the hardware installation necessary for standalone Oracle Exadata Storage Servers connecting to Oracle Exadata Database Machine.

1. Unpack, situate, and set up the storage server.
2. Read the printed vendor documentation supplied with the storage servers.
3. Connect the management and Integrated Lights Out Manager (ILOM) Ethernet ports to the management network.
Both the Oracle Exadata Database Machine and additional Oracle Exadata Storage Servers must use the same type of RDMA Network Fabric.

5. Attach a computer or terminal to one of the new storage servers so you can configure the cell as described in Configuring Oracle Exadata System Software for Your Location. You must configure each new cell.

6. Connect the power supply on the new storage servers.

7. Power on the storage servers.

8. Assign IP addresses for the network interconnections of the new storage servers.

Related Topics

• Assigning IP Addresses for Oracle Exadata Database Machine
  This topic summarizes the Oracle Exadata Database Machine network preparation before installing the new storage server.

B.1.2 Installing the Oracle Exadata System Software on a Cell

This procedure describes how to install the Oracle Exadata System Software on an Oracle Exadata Storage Server.

1. Power on the Oracle Exadata Storage Server to boot its operating system.

2. Log in to the Oracle Exadata Storage Server with super user (root) privileges.

3. Change the root password on the cell.

4. Verify that the Oracle Exadata Storage Server meets the software requirements.

5. Copy the self-extracting binary to the Oracle Exadata Storage Server, if necessary. You can use the ftp command to copy the binary to the Oracle Exadata Storage Server from a network location, or you can copy the binary from a portable storage device.

6. Install the self-extracting binary on the Oracle Exadata Storage Server if it is not already installed.

   ./cell_package_version.bin

7. Review the /opt/oracle/cell/.install_log.txt file to verify that the installation completed correctly. The installation process adds two new users, celladmin and cellmonitor. The celladmin user should be used to run all services on the cell. The cellmonitor user is for monitoring purposes.
After successful installation of the Oracle Exadata System Software, follow the instructions at "Preparing the Servers".

Related Topics
- Preparing the Servers
  Use the following steps to prepare the database servers and storage servers for use.

B.1.3 Updates Made During Installation

Changes that are made to snmpd files during installation include the following:

- Updates to the public community string
- New entries for trapdestination
- Updates to the kernel configuration files (limits for aio, udprecv, and number of open files)

B.1.4 Online Documentation

The documentation can be found online at https://docs.oracle.com/.

B.2 Background Processes in the Oracle Exadata Storage Server Environment

The background processes for the database and Oracle ASM instance for an Oracle Exadata Storage Server environment are the same as other environments, except for the following background processes:

- diskmon Process
- XDMG Process
- XDWK Process

B.2.1 diskmon Process

The diskmon process is a fundamental component of Oracle Exadata System Software, and is responsible for implementing I/O fencing. The process is located on the database server host computer, and is part of Oracle Clusterware Cluster Ready Services (CRS). This process is important for Oracle Exadata System Software and should not be modified.

The log files for diskmon are located in the $CRS_HOME/log/hostname/diskmon directory.
B.2.2 XDMG Process

The XDMG (Exadata Automation Manager) process initiates automation tasks used for monitoring storage. This background process monitors all configured Oracle Exadata Storage Servers for state changes, such as replaced disks, and performs the required tasks for such changes. Its primary task is to watch for inaccessible disks and cells, and to detect when the disks and cells become accessible. When the disks and cells are accessible, the XDMG process initiates the ASM ONLINE process, which is handled by the XDWK background process. The XDMG process runs in the Oracle ASM instances.

B.2.3 XDWK Process

The XDWK (Exadata Automation Worker) process performs automation tasks requested by the XDMG background process. The XDWK process begins when asynchronous actions, such as ONLINE, DROP or ADD for an Oracle ASM disk are requested by the XDMG process. The XDWK process stops after 5 minutes of inactivity. The XDWK process runs in the Oracle ASM instances.
C

Exadata-Specific Information in Oracle Database Dictionary Views

- Oracle Database Dictionary Views
- Automatic Workload Repository Views

C.1 Oracle Database Dictionary Views

The Oracle Database dictionary contains views that are specific to Oracle Exadata Database Machine.

- Using the V$CELL and GV$CELL Views to Display Oracle Exadata Storage Server Identification
  The V$CELL view provides identifying information about cells.

- About the V$CELL_ Views
  The views prefixed with V$CELL_ contain internal statistical information, which is used by Oracle Support Services.

- Using V$BACKUP_DATAFILE with Oracle Exadata Storage Server
  The V$BACKUP_DATAFILE view contains columns relevant to Oracle Exadata Storage Server during Oracle Recovery Manager (RMAN) incremental backups.

- Using V$ASM_DISK_SPARSE and V$ASM_DISKGROUP_SPARSE to Monitor Sparse Disks
  The V$ASM_DISK_SPARSE and V$ASM_DISKGROUP_SPARSE views contain information about sparse disks.

C.1.1 Using the V$CELL and GV$CELL Views to Display Oracle Exadata Storage Server Identification

The V$CELL view provides identifying information about cells.

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELL_HASHVAL</td>
<td>A numeric hash value for the cell. For example: 138889696</td>
</tr>
</tbody>
</table>

**Note:** The cell hash value is often contained in the P1 argument for cell-related wait events in the V$SESSION_WAIT and V$ACTIVE_SESSION_HISTORY views. You can join to this column to identify the corresponding cell.
Table C-1  (Cont.) V$CELL View Columns and Descriptions

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELL_PATH</td>
<td>A character string (maximum 400) that specifies the IP addresses of the cell. These are the IP addresses specified in the cellip.ora file.</td>
</tr>
<tr>
<td>CELL_TYPE</td>
<td>The type of storage cell. Either EXADATA or BDSQL.</td>
</tr>
</tbody>
</table>

The GV$CELL view contains the same columns as the V$CELL view, and includes the INST_ID column. The INST_ID column displays the instance number from which the associated V$ view information was obtained. Querying a GV$ view retrieves the V$ view information from all qualified instances.

C.1.2 About the V$CELL_ Views

The views prefixed with V$CELL_ contain internal statistical information, which is used by Oracle Support Services.

This includes the following views:

- V$CELL_CONFIG
- V$CELL_CONFIG_INFO
- V$CELL_THREAD_HISTORY
- V$CELL_DB
- V$CELL_DISK
- V$CELL_DISK_HISTORY
- V$CELL_GLOBAL
- V$CELL_GLOBAL_HISTORY
- V$CELL_IOREASON
- V$CELL_IOREASON_NAME
- V$CELL_METRIC_DESC
- V$CELL_REQUEST_TOTALS
- V$CELL_STATE
- V$CELL_THREAD_HISTORY
- V$CELL_OFL_THREAD_HISTORY
- V$CELL_OPEN_ALERTS

These views contain instantaneous statistical information that may be easily misinterpreted or incomplete when viewed in isolation. Instead, use the corresponding Automatic Workload Repository (AWR) views, which are prefixed with DBA_HIST_CELL_. Or, use the AWR report.

Related Topics

- Automatic Workload Repository Views
C.1.3 Using V$BACKUP_DATAFILE with Oracle Exadata Storage Server

The V$BACKUP_DATAFILE view contains columns relevant to Oracle Exadata Storage Server during Oracle Recovery Manager (RMAN) incremental backups.

Table C-2 V$BACKUP_DATAFILE Columns and Descriptions

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLOCKS</td>
<td>Size of the backup data file in blocks.</td>
</tr>
<tr>
<td>BLOCKS_READ</td>
<td>The number of blocks that were scanned while taking this backup. If this is</td>
</tr>
<tr>
<td></td>
<td>an incremental backup, and block change tracking was used to optimize the</td>
</tr>
<tr>
<td></td>
<td>backup, then the value of this column is smaller than DATAFILE_BLOCKS.</td>
</tr>
<tr>
<td></td>
<td>Otherwise, the value of this column is the same as DATAFILE_BLOCKS.</td>
</tr>
<tr>
<td>BLOCKS_SKIPPED_IN_CELL</td>
<td>The number of blocks that were read and filtered at the Oracle Exadata</td>
</tr>
<tr>
<td></td>
<td>Storage Server to optimize the RMAN incremental backup.</td>
</tr>
<tr>
<td>DATAFILE_BLOCKS</td>
<td>Size of the data file in blocks at backup time. This value is also the</td>
</tr>
<tr>
<td></td>
<td>number of blocks taken by the data file restarted from this backup.</td>
</tr>
</tbody>
</table>

The percentage of blocks skipped by Oracle Exadata System Software is calculated as follows:

\[
\frac{\text{BLOCKS_SKIPPED_IN_CELL}}{\text{BLOCKS_READ}} \times 100
\]

This number changes significantly based on block change tracking.

If block change tracking for fast incremental backups is used, then most of the filtering is done at the database using the change tracking file, and the blocks are skipped before making an I/O request to the cell. If block change tracking is not used, then all of the blocks are filtered at the cell.

Related Topics
- Oracle Database Reference
- Oracle Database Backup and Recovery User’s Guide

C.1.4 Using V$ASM_DISK_SPARSE and V$ASM_DISKGROUP_SPARSE to Monitor Sparse Disks

The V$ASM_DISK_SPARSE and V$ASM_DISKGROUP_SPARSE views contain information about sparse disks.

Table C-3 V$ASM_DISK_SPARSE Columns and Descriptions

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP_NUMBER</td>
<td>The number of the disk group containing the disk.</td>
</tr>
</tbody>
</table>
Table C-3  (Cont.) V$ASM_DISK_SPARSE Columns and Descriptions

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISK_NUMBER</td>
<td>The number assigned to the disk within this disk group.</td>
</tr>
<tr>
<td>INCARNATION</td>
<td>The incarnation number for the disk.</td>
</tr>
<tr>
<td>ALLOCATED_MAT_MB</td>
<td>The total used physical and materialized capacity on the disk.</td>
</tr>
<tr>
<td>TOTAL_MAT_MB</td>
<td>The total physical capacity on the disk.</td>
</tr>
<tr>
<td>SPARSE_READS</td>
<td>The total number of I/O read requests on non-materialized regions of the disk.</td>
</tr>
<tr>
<td>SPARSE_BYTES_READ</td>
<td>The total number of bytes read from non-materialized regions of the disk.</td>
</tr>
<tr>
<td>SPARSE_READ_TIME</td>
<td>The time taken by sparse read I/O operations.</td>
</tr>
</tbody>
</table>

Table C-4  V$ASM_DISKGROUP_SPARSE Columns and Descriptions

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP_NUMBER</td>
<td>The cluster-wide number assigned to the disk group.</td>
</tr>
<tr>
<td>ALLOCATED_MAT_MB</td>
<td>The total used physical and materialized capacity of the disk group.</td>
</tr>
<tr>
<td>TOTAL_MAT_MB</td>
<td>The total physical capacity of the disk group.</td>
</tr>
</tbody>
</table>

C.2 Automatic Workload Repository Views

The Automatic Workload Repository (AWR) views contain information that is specific to Oracle Exadata Database Machine.

- `DBA_HIST_ASM_BAD_DISK`
- `DBA_HIST_ASM_DISKGROUP`
- `DBA_HIST_ASM_DISKGROUP_STAT`
- `DBA_HIST_ASM_DISK_STAT_SUMMARY`
- `DBA_HIST_CELL_CONFIG`
- `DBA_HIST_CELL_CONFIG_DETAIL`
- `DBA_HIST_CELL_DB`
- `DBA_HIST_CELL_DISKTYPE`
- `DBA_HIST_CELL_DISK_NAME`
- `DBA_HIST_CELL_DISK_SUMMARY`
- `DBA_HIST_CELL_GLOBAL`
- `DBA_HIST_CELL_GLOBAL_SUMMARY`
- `DBA_HIST_CELL_IOREASON`
- `DBA_HIST_CELL_IOREASON_NAME`
- `DBA_HIST_CELL_METRIC_DESC`
The `DBA_HIST_ASM_BAD_DISK` view displays historic information about non-online Oracle Automatic Storage Management (Oracle ASM) disks. This view contains snapshots of `V$ASM_DISK`.

### C.2.1 DBA_HIST_ASM_BAD_DISK

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>GROUP_NUMBER</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Number of the disk group containing the disk.</td>
</tr>
<tr>
<td>NAME</td>
<td>VARCHAR2(128)</td>
<td>NOT NULL</td>
<td>Name of the disk.</td>
</tr>
<tr>
<td>PATH</td>
<td>VARCHAR2(256)</td>
<td>NOT NULL</td>
<td>Operating system path name portion of the name returned by discovery.</td>
</tr>
<tr>
<td>STATUS</td>
<td>VARCHAR2(8)</td>
<td>NOT NULL</td>
<td>Global status for the disk. Only non-online disks are stored in this view.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

The `DBA_HIST_ASM_DISKGROUP` view contains information about Oracle ASM disk groups. This retrieves a subset of columns from `V$ASM_DISKGROUP`.

### C.2.2 DBA_HIST_ASM_DISKGROUP

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CURRENT_SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>AWR snapshot identifier for when the disk group was captured.</td>
</tr>
<tr>
<td>GROUP_NUMBER</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Cluster-wide number assigned to the disk group.</td>
</tr>
<tr>
<td>NAME</td>
<td>VARCHAR2(128)</td>
<td>NOT NULL</td>
<td>Name of the disk group.</td>
</tr>
<tr>
<td>TYPE</td>
<td>VARCHAR2(6)</td>
<td>NOT NULL</td>
<td>Redundancy type for the disk group.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
</tbody>
</table>
### C.2.3 DBA_HIST_ASM_DISKGROUP_STAT

The **DBA_HIST_ASM_DISKGROUP_STAT** view displays historic information about Oracle ASM disk groups.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>GROUP_NUMBER</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Cluster-wide number assigned to the disk group.</td>
</tr>
<tr>
<td>TOTAL_MB</td>
<td>NUMBER</td>
<td></td>
<td>Total capacity of the disk group in MB.</td>
</tr>
<tr>
<td>FREE_MB</td>
<td>NUMBER</td>
<td></td>
<td>Unused capacity in MB of the disk group.</td>
</tr>
<tr>
<td>NUM_DISK</td>
<td>NUMBER</td>
<td></td>
<td>Number of disks in the disk group.</td>
</tr>
<tr>
<td>NUM_FAILGROUP</td>
<td>NUMBER</td>
<td></td>
<td>Number of failure groups in the disk group.</td>
</tr>
<tr>
<td>STATE</td>
<td>VARCHAR2(32)</td>
<td></td>
<td>State of the disk group.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

### C.2.4 DBA_HIST_ASM_DISK_STAT_SUMMARY

The **DBA_HIST_ASM_DISK_STAT_SUMMARY** view displays historic summary information about Oracle ASM disk groups. It aggregates information from **V$ASM_DISK_STAT**.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>INSTANCE_NUMBE R</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database instance number.</td>
</tr>
</tbody>
</table>
### C.2.5 DBA_HIST_CELL_CONFIG

The DBA_HIST_CELL_CONFIG view contains information about the configuration of the cells. This data is from V$CELL_CONFIG_INFO.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CURRENT_SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>AWR snapshot identifier for when the configuration was captured.</td>
</tr>
<tr>
<td>CELLNAME</td>
<td>VARCHAR2(256)</td>
<td>NOT NULL</td>
<td>Unique identifier for the cell.</td>
</tr>
<tr>
<td>Column</td>
<td>Datatype</td>
<td>NULL</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CELLHASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number to uniquely identify the cell.</td>
</tr>
<tr>
<td>CONTYPE</td>
<td>VARCHAR2(15)</td>
<td>NOT NULL</td>
<td>Configuration type. The configuration type determines the type of information stored in the CONFVAL column.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- AWRXML: configuration information pertinent to AWR reports.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- CELL: configuration information about the cell.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- CELLDISK: configuration information about cell disks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- GRIDDISK: configuration information about grid disks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- IORM_OBJ: IORM objective.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- OFFLOAD: configuration information for offload servers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- PHYSICALDISK: configuration information about physical disks.</td>
</tr>
<tr>
<td>CONFVAL</td>
<td>CLOB</td>
<td></td>
<td>XML data associated with CONTYPE.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

C.2.6 DBA_HIST_CELL_CONFIG_DETAIL

The DBA_HIST_CELL_CONFIG_DETAIL view displays historic information about the configuration of the cells. This view has the same descriptions as DBA_HIST_CELL_CONFIG, but includes the data for each snapshot, such as a historic view.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELLNAME</td>
<td>VARCHAR2(256)</td>
<td>NOT NULL</td>
<td>Unique identifier for the cell.</td>
</tr>
<tr>
<td>CELLHASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number to uniquely identify the cell.</td>
</tr>
<tr>
<td>Column</td>
<td>Datatype</td>
<td>NULL</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CONFTYPE</td>
<td>VARCHAR2(15)</td>
<td>NOT NULL</td>
<td>Configuration type. The configuration type determines the type of information stored in the CONFVAL column.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- AWRXML: configuration information pertinent to AWR reports.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- CELL: configuration information about the cell.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- CELLDISK: configuration information about cell disks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- GRIDDISK: configuration information about grid disks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- IORM_OBJ: IORM objective.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- OFFLOAD: configuration information for offload servers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- PHYSICALDISK: configuration information about physical disks.</td>
</tr>
<tr>
<td>CONFVAL</td>
<td>CLOB</td>
<td></td>
<td>XML data associated with the CONFTYPE.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

C.2.7 DBA_HIST_CELL_DB

The DBA_HIST_CELL_DB view displays historic information about the databases consuming resources on the cells. Only the top 10 databases are stored in each AWR snapshot. This is a snapshot of V$CELL_DB.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number to uniquely identify the cell.</td>
</tr>
<tr>
<td>INCARNATION_NUM</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Incarnation number of the cell. Each cell reboot will increment the incarnation number.</td>
</tr>
<tr>
<td>SRC_DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier performing the I/Os.</td>
</tr>
<tr>
<td>SRC_DBNAME</td>
<td>VARCHAR2(256)</td>
<td></td>
<td>DB Name performing the I/Os.</td>
</tr>
<tr>
<td>DISK_REQUEST</td>
<td>NUMBER</td>
<td></td>
<td>Number of disk I/O requests performed by the database.</td>
</tr>
<tr>
<td>Column</td>
<td>Datatype</td>
<td>NULL</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------</td>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DISK_BYTES</td>
<td>NUMBER</td>
<td></td>
<td>Number of disk I/O bytes processed by the database.</td>
</tr>
<tr>
<td>FLASH_REQUETS</td>
<td>NUMBER</td>
<td></td>
<td>Number of flash I/O requests performed by the database.</td>
</tr>
<tr>
<td>FLASH_BYTES</td>
<td>NUMBER</td>
<td></td>
<td>Number of flash I/O bytes processed by the database.</td>
</tr>
<tr>
<td>DISK_SMALL_IO_REQS</td>
<td>NUMBER</td>
<td></td>
<td>Number of small IO requests issued to disks by the database.</td>
</tr>
<tr>
<td>DISK_LARGE_IO_REQS</td>
<td>NUMBER</td>
<td></td>
<td>Number of large IO requests issued to disks by the database.</td>
</tr>
<tr>
<td>FLASH_SMALL_IO_REQS</td>
<td>NUMBER</td>
<td></td>
<td>Number of small IO requests issued to flash by the database.</td>
</tr>
<tr>
<td>FLASH_LARGE_IO_REQS</td>
<td>NUMBER</td>
<td></td>
<td>Number of large IO requests issued to flash by the database.</td>
</tr>
<tr>
<td>DISK_SMALL_IO_SERVICE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The latency (in nanoseconds) of small IO requests issued to disks by the database.</td>
</tr>
<tr>
<td>DISK_LARGE_IO_SERVICE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The latency (in nanoseconds) of large IO requests issued to disks by the database.</td>
</tr>
<tr>
<td>DISK_SMALL_IO_QUEUE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The IORM wait time (in nanoseconds) for small IO requests issued to disks by the database.</td>
</tr>
<tr>
<td>DISK_LARGE_IO_QUEUE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The IORM wait time (in nanoseconds) for large IO requests issued to disks by the database.</td>
</tr>
<tr>
<td>FLASH_SMALL_IO_SERVICE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The latency (in nanoseconds) of small IO requests issued to flash by the database.</td>
</tr>
<tr>
<td>FLASH_SMALL_IO_QUEUE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The IORM wait time (in nanoseconds) for small IO requests issued to flash by the database.</td>
</tr>
<tr>
<td>FLASH_LARGE_IO_SERVICE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The latency (in nanoseconds) of large IO requests issued to flash by the database.</td>
</tr>
<tr>
<td>FLASH_LARGE_IO_QUEUE_TIME</td>
<td>NUMBER</td>
<td></td>
<td>The IORM wait time (in nanoseconds) for large IO requests issued to flash by the database.</td>
</tr>
</tbody>
</table>
C.2.8 DBA_HIST_CELL_DISKTYPE

The DBA_HIST_CELL_DISKTYPE view displays historic information about cells, including the types of disks and the capacity of the disks. This view is derived from DBA_HIST_CELL_CONFIG and DBA_HIST_CELL_CONFIG_DETAIL.
### C.2.9 DBA_HIST_CELL_DISK_NAME

The **DBA_HIST_CELL_DISK_NAME** view displays historic information about disk names on cells. This view is derived from **DBA_HIST_CELL_CONFIG** and **DBA_HIST_CELL_CONFIG_DETAIL**.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number to uniquely identify the cell.</td>
</tr>
<tr>
<td>DISK_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique identifier of the disk.</td>
</tr>
<tr>
<td>DISK_NAME</td>
<td>VARCHAR2(4000)</td>
<td></td>
<td>User-readable cell name.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX_DISK_MBP</td>
<td>NUMBER</td>
<td></td>
<td>The maximum I/O throughput for the hard disk type, in megabytes per second.</td>
</tr>
<tr>
<td>MAX_FLASH_MB</td>
<td>NUMBER</td>
<td></td>
<td>The maximum I/O throughput for the flash disk type, in megabytes per second.</td>
</tr>
<tr>
<td>MAX_CELL_DIS</td>
<td>NUMBER</td>
<td></td>
<td>The maximum number of hard disk IOPs for the cell. This is calculated using</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MAX_DISK_IOPS*NUM_HARD_DISKS.</td>
</tr>
<tr>
<td>MAX_CELL_FLASH</td>
<td>NUMBER</td>
<td></td>
<td>The maximum number of flash IOPs for the cell. This is calculated using</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MAX_FLASH_IOPS*NUM_FLASH_DISKS.</td>
</tr>
<tr>
<td>MAX_CELL_DISK_MBPS</td>
<td>NUMBER</td>
<td></td>
<td>The maximum I/O throughput for the cell, in megabytes per second. This is calculated using MAX_DISK_MBPS*NUM_HARD_DISKS.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number to uniquely identify the cell.</td>
</tr>
<tr>
<td>DISK_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique identifier of the disk.</td>
</tr>
<tr>
<td>DISK_NAME</td>
<td>VARCHAR2(4000)</td>
<td></td>
<td>User-readable cell name.</td>
</tr>
</tbody>
</table>
The `DBA_HIST_CELL_DISK_SUMMARY` view displays historic information about the performance of disks on cells. For each AWR snapshot, a summary of the per-minute metrics is visible in this view. Each metric stores the sum and the sum of squares, similar to `DBA_HIST_SYSMETRIC_SUMMARY`. This view is a summary of `V$CELL_DISK_HISTORY`.

### Column Details

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISK</td>
<td>VARCHAR2(4000)</td>
<td>NULL</td>
<td>Disk type, either FlashDisk or HardDisk.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

### Summary of Columns

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number to uniquely identify the cell.</td>
</tr>
<tr>
<td>DISK_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique identifier of the disk.</td>
</tr>
<tr>
<td>NUM_SAMPLES</td>
<td>NUMBER</td>
<td></td>
<td>Number of one-minute samples the summary is based on.</td>
</tr>
<tr>
<td>DISK_UTILIZATION_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute disk utilization metrics. This is based on OS statistics.</td>
</tr>
<tr>
<td>READS_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute read requests metrics per second. This is based on OS statistics.</td>
</tr>
<tr>
<td>WRITE_MB_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute write metrics, in megabytes per second. This is based on OS statistics.</td>
</tr>
<tr>
<td>WRITE_REQUESTS_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute IOPs. This is based on OS statistics.</td>
</tr>
<tr>
<td>IO_MB_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute I/O metrics, in megabytes per second. This is based on OS statistics.</td>
</tr>
<tr>
<td>Column</td>
<td>Datatype</td>
<td>NULL</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------</td>
<td>------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SERVICE_TIME_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute service time metrics. This is based on OS statistics.</td>
</tr>
<tr>
<td>WAIT_TIME_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute wait time metrics. This is based on OS statistics.</td>
</tr>
<tr>
<td>SMALL_READS_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute small read requests metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>SMALL_WRITES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute small write requests metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>LARGE_READS_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute large read requests metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>LARGE_WRITES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute large write requests metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>SMALL_READ_BYTES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute small read bytes metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>SMALL_WRITE_BYTES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute small write bytes metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>LARGE_READ_BYTES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute large read bytes metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>LARGE_WRITE_BYTES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute large write bytes metrics per second. This is from cell server statistics.</td>
</tr>
<tr>
<td>SMALL_READ_LATENCY_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute small read latency metrics. This data is from cell server statistics.</td>
</tr>
<tr>
<td>SMALL_WRITE_LATENCY_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute small write latency metrics. This data is from cell server statistics.</td>
</tr>
<tr>
<td>LARGE_READ_LATENCY_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute large read latency metrics. This data is from cell server statistics.</td>
</tr>
<tr>
<td>LARGE_WRITE_LATENCY_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute large write latency metrics. This data is from cell server statistics.</td>
</tr>
<tr>
<td>APP_IO_REQUESTS_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute IOPs metrics. This data is from cell server statistics.</td>
</tr>
<tr>
<td>APP_IO_BYTES_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute I/O metrics, in megabytes per second. This data is from cell server statistics.</td>
</tr>
</tbody>
</table>
### C.2.11 DBA_HIST_CELL_GLOBAL

The DBA_HIST_CELL_GLOBAL view displays historic information about cell performance statistics. This is a snapshot of V$CELL_GLOBAL.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number that uniquely identifies the cell.</td>
</tr>
<tr>
<td>INCARNATION_NUM</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Incarnation number of the cell.</td>
</tr>
<tr>
<td>METRIC_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Metric_id that identifies the metric.</td>
</tr>
<tr>
<td>METRIC_NAME</td>
<td>VARCHAR2(257)</td>
<td></td>
<td>Name of the metric.</td>
</tr>
<tr>
<td>METRIC_VALUE</td>
<td>NUMBER</td>
<td></td>
<td>The value of the metric.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>
The `DBA_HIST_CELL_GLOBAL_SUMMARY` view displays historic information about the performance of cells. For each AWR snapshot, a summary of the per-minute metrics is visible in this view. This is a summary of `V$CELL_GLOBAL_HISTORY`.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot ID</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number that uniquely identifies the cell.</td>
</tr>
<tr>
<td>INCARNATION_NUM</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Incarnation number of the cell.</td>
</tr>
<tr>
<td>NUM_SAMPLES</td>
<td>NUMBER</td>
<td></td>
<td>Number of 1 minute samples the summary is based on.</td>
</tr>
<tr>
<td>CPU_USAGE_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute CPU usage percentage metrics.</td>
</tr>
<tr>
<td>SYS_USAGE_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute system CPU usage percentage metrics.</td>
</tr>
<tr>
<td>USER_USAGE_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute user CPU usage percentage metrics.</td>
</tr>
<tr>
<td>NETWORK_BYTE_S_RECD_SUM</td>
<td>NUMBER</td>
<td></td>
<td>Sum of the per-minute network received metrics, in bytes per second.</td>
</tr>
<tr>
<td>*_SUMX2</td>
<td>NUMBER</td>
<td></td>
<td>All previous columns from CPU_USAGE to NETWORKBYTES_SENT are repeated with the SUMX2 suffix. This is the sum of squares of the per-minute metrics.</td>
</tr>
<tr>
<td>*_AVG</td>
<td>NUMBER</td>
<td></td>
<td>All previous columns from CPU_USAGE to NETWORKBYTES_SENT are repeated with the SUMX2 suffix. This is the average of the per-minute metrics.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>
C.2.13 DBA_HIST_CELL_IOREASON

The DBA_HIST_CELL_IOREASON view displays historic information about the reasons for performing I/Os on the cells. This is a snapshot of V$CELL_IOREASON.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number that uniquely identifies the cell.</td>
</tr>
<tr>
<td>INCARNATION_</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Incarnation number of the cell.</td>
</tr>
<tr>
<td>NUM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REASON_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique identifier for the I/O reason.</td>
</tr>
<tr>
<td>REASON_NAME</td>
<td>VARCHAR2(257)</td>
<td></td>
<td>User readable I/O reason.</td>
</tr>
<tr>
<td>REQUESTS</td>
<td>NUMBER</td>
<td></td>
<td>Number of I/O requests performed for the reason.</td>
</tr>
<tr>
<td>BYTES</td>
<td>NUMBER</td>
<td></td>
<td>Number of IO Bytes performed for the reason.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

C.2.14 DBA_HIST_CELL_IOREASON_NAME

The DBA_HIST_CELL_IOREASON_NAME view displays the reasons for different I/Os. This data is from V$CELL_IOREASON_NAME.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>REASON_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique identifier for an I/O reason.</td>
</tr>
<tr>
<td>REASON_NAME</td>
<td>VARCHAR2(257)</td>
<td></td>
<td>User-readable description of the I/O reason.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>
C.2.15 DBA_HIST_CELL_METRIC_DESC

The DBA_HIST_CELL_METRIC_DESC view displays information about cell metrics. This data is from V$CELL_METRIC_DESC.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>METRIC_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique identifier for the metric</td>
</tr>
<tr>
<td>METRIC_NAME</td>
<td>VARCHAR2(257)</td>
<td></td>
<td>User-readable metric name</td>
</tr>
<tr>
<td>METRIC_TYPE</td>
<td>VARCHAR2(17)</td>
<td>NOT NULL</td>
<td>Unit for the metrics, if NULL this is a count.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

C.2.16 DBA_HIST_CELL_NAME

The DBA_HIST_CELL_NAME view displays historic information about the names of cells. This view is derived from DBA_HIST_CELL_CONFIG and DBA_HIST_CELL_CONFIG_DETAIL.

<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number that uniquely identifies the cell.</td>
</tr>
<tr>
<td>CELL_NAME</td>
<td>VARCHAR2(4000)</td>
<td></td>
<td>User-readable cell name.</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>

C.2.17 DBA_HIST_CELL_OPEN_ALERTS

The DBA_HIST_CELL_OPEN_ALERTS view displays the open alerts on cells. Only the 10 most-recent alerts for each cell are stored in AWR. This data is from V$CELL_OPEN_ALERTS.
<table>
<thead>
<tr>
<th>Column</th>
<th>Datatype</th>
<th>NULL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP_ID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Unique snapshot identifier.</td>
</tr>
<tr>
<td>DBID</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Database identifier for the snapshot.</td>
</tr>
<tr>
<td>CELL_HASH</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Hash number that uniquely identifies the cell.</td>
</tr>
<tr>
<td>BEGIN_TIME</td>
<td>DATE</td>
<td>NOT NULL</td>
<td>Time that the alert was opened.</td>
</tr>
<tr>
<td>SEQ_NO</td>
<td>NUMBER</td>
<td>NOT NULL</td>
<td>Sequence number of the alert.</td>
</tr>
<tr>
<td>MESSAGE</td>
<td>VARCHAR2(1024)</td>
<td>NOT NULL</td>
<td>Message describing the alert.</td>
</tr>
<tr>
<td>STATEFUL</td>
<td>VARCHAR2(1)</td>
<td></td>
<td>Whether or not the alert has state.</td>
</tr>
<tr>
<td>SEVERITY</td>
<td>VARCHAR2(64)</td>
<td></td>
<td>The severity of the alert. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• critical</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• info</td>
</tr>
<tr>
<td>CON_DBID</td>
<td>NUMBER</td>
<td></td>
<td>The database identifier.</td>
</tr>
<tr>
<td>CON_ID</td>
<td>NUMBER</td>
<td></td>
<td>The identifier of the container identified by CON_DBID. Possible values are as follows:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 0: This value is used for non-container database (CDB).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1: This value is used for the root container of the CDB.</td>
</tr>
</tbody>
</table>
Use these tips for using screen readers and screen magnifiers with Oracle Exadata System Software utilities.

Oracle Exadata System Software includes tools such as dcli and cellcli that you can run from the command line.

- **Tips on Using Screen Readers and Braille Displays**
  Use the following tips when using screen readers with Oracle Exadata System Software

- **Tips on Using Screen Magnifiers**
  Use the following tips when using screen magnifiers with Oracle Exadata System Software

- **Tips on Using Exawatcher Charts**
  Exawatcher is an Exadata specific tool that collects performance data from Exadata storage cells.

- **Oracle Exadata Deployment Assistant (OEDA) Web interface Accessibility**
  The OEDA Web interface conforms with the Web Content Accessibility Guidelines version 2.0 at the AA level (WCAG 2.0 AA).

## D.1 Tips on Using Screen Readers and Braille Displays

Use the following tips when using screen readers with Oracle Exadata System Software.

Examples of screen readers include JAWS, SuperNova, NVDA, and Window-Eyes. Each of these provides text-to-speech output and supports braille displays.

### Note:

Window-Eyes is now a part of ZoomText Fusion by Ai Squared.

- Use a character mode based terminal such as Putty or Cygwin. Do not use an X-Windows-based VNC.

- For screen reader users, we recommend installing "screen" in order to get multiple session support. The Linux based screen program allows for multiple sessions in different windows. You can access each session with keyboard based commands, for example, `Ctrl-a`. Screen allows you to detach or re-attach to a given window session. Like VNC, if you get disconnected when running Oracle ExaCHK, or patchmgr, or other program, you can re-attach to and resume that session.
The screen package is not installed by default on Exadata. You will need to install it using `yum`. See the "How To Use Linux Screen" tutorial at https://www.rackaid.com/blog/linux-screen-tutorial-and-how-to/ for details.

- In the settings of the terminal software, set the cursor type to "block" cursor, not blinking or flashing.
- The output of the commands can generate a significant amount of information and might spill off the terminal window, and the virtual window or braille display. For example, the following command can generate a long alert history output:

```
dcli -g cell_group -l root cellcli list alerthistory
```

To display the output one screen-full at a time, pipe the output through the `more` command, as in the following:

```
dcli -g cell_group -l root cellcli list alerthistory | more
```

You can then use the space bar key to page through the output.

- When `exachk` or `dbnodeupdate.sh` is launched interactively, do not pipe its output to the `more` or `page` commands. As it runs, it displays informational messages on the terminal. The messages pause when `exachk` requires user input, then resume after input is received. Important messages, user input, errors, and check results are logged in various files. The results from `exachk` are written to an HTML report. All you need to do is to transfer the HTML report to a computer that runs your assistive technology and open the HTML report in a browser that you can access with your assistive technology.

- If you are running the `patchmgr` utility, and it is performing a task that takes some time to complete, the output displays a "spinner" and a countdown clock. The "spinner" cycles through the \, |, and / characters in-place, and the countdown clock is updated periodically. When the task is done, the output displays a "success" or "error" message, depending on the outcome. The output messages are also logged in a log file.

- A few recommended screen reader settings include the following (JAWS is used here just as an example):
  - Set the JAWS cursor to "All". Use the key combination of Insert + s until you hear "All".
  - You may need to turn off virtual cursor. If you are using JAWS, you can do this using the key combination of Insert + z.
  - Use the virtual window to capture text. If you are using JAWS, you can do this using the key combination of Insert + Alt + w.

### D.2 Tips on Using Screen Magnifiers

Use the following tips when using screen magnifiers with Oracle Exadata System Software

Examples of screen magnifiers include ZoomText, MAGic, and SuperNova.

- Screen magnifiers can support both character-based terminals and X-Window-based VNC.
D.3 Tips on Using Exawatcher Charts

Exawatcher is an Exadata specific tool that collects performance data from Exadata storage cells.

The data collected by Exawatcher can be graphed and presented in web pages. The web pages are based on Oracle JavaScript Extension Toolkit (JET), HTML and some use of ARIA. The Exawatcher pages can present charts which can be navigated using only a keyboard.

D.4 Oracle Exadata Deployment Assistant (OEDA) Web interface Accessibility

The OEDA Web interface conforms with the Web Content Accessibility Guidelines version 2.0 at the AA level (WCAG 2.0 AA).

The OEDA Web interface was built using Oracle JET, which provides support for:

- **Keyboard and touch navigation**

- **Zoom**
  Oracle JET supports browser zooming up to 200%. For example, on the Firefox browser, you can choose **View**, then **Zoom**, and then **Zoom In**.

- **Screen reader**
  Oracle JET supports screen readers such as JAWS, Apple VoiceOver, and Google Talkbalk by generating content that complies with WAI-ARIA standards, and no special mode is needed.

- **Oracle JET component roles and names**
  Each component has an appropriate role, such as `button`, `link`, and so on, and each component supports an associated name (label), if applicable.

- **Sufficient color contrast**
  Oracle JET provides the Alta theme which is designed to provide a luminosity contrast ratio of at least 4.5:1.
Related Topics

• WAI-ARIA: Developing a Keyboard Interface