

Oracle® ZFS Storage Appliance 管理指南



文件号码 E54240-02
2014 年 6 月

版权所有 © 2009, 2014, Oracle 和/或其附属公司。保留所有权利。

本软件和相关文档是根据许可证协议提供的，该许可证协议中规定了关于使用和公开本软件和相关文档的各种限制，并受知识产权法的保护。除非在许可证协议中明确许可或适用法律明确授权，否则不得以任何形式、任何方式使用、拷贝、复制、翻译、广播、修改、授权、传播、分发、展示、执行、发布或显示本软件和相关文档的任何部分。除非法律要求实现互操作，否则严禁对本软件进行逆向工程设计、反汇编或反编译。

此文档所含信息可能随时被修改，恕不另行通知，我们不保证该信息没有错误。如果贵方发现任何问题，请书面通知我们。

如果将本软件或相关文档交付给美国政府，或者交付给以美国政府名义获得许可证的任何机构，必须符合以下规定：

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

本软件或硬件是为了在各种信息管理应用领域内的一般使用而开发的。它不应被应用于任何存在危险或潜在危险的应用领域，也不是为此而开发的，其中包括可能会产生人身伤害的应用领域。如果在危险应用领域内使用本软件或硬件，贵方应负责采取所有适当的防范措施，包括备份、冗余和其它确保安全使用本软件或硬件的措施。对于因在危险应用领域内使用本软件或硬件所造成的一切损失或损害，Oracle Corporation 及其附属公司概不负责。

Oracle 和 Java 是 Oracle 和/或其附属公司的注册商标。其他名称可能是各自所有者的商标。

Intel 和 Intel Xeon 是 Intel Corporation 的商标或注册商标。所有 SPARC 商标均是 SPARC International, Inc 的商标或注册商标，并应按照许可证的规定使用。AMD、Opteron、AMD 徽标以及 AMD Opteron 徽标是 Advanced Micro Devices 的商标或注册商标。UNIX 是 The Open Group 的注册商标。

本软件或硬件以及文档可能提供了访问第三方内容、产品和服务的方式或有关这些内容、产品和服务的信息。对于第三方内容、产品和服务，Oracle Corporation 及其附属公司明确表示不承担任何种类的担保，亦不对其承担任何责任。对于因访问或使用第三方内容、产品或服务所造成的任何损失、成本或损害，Oracle Corporation 及其附属公司概不负责。

目录

使用本文档	19
1 Oracle ZFS Storage Appliance 概述	21
ZFSSA 主要功能	21
支持的协议	21
ZFSSA 数据服务	22
数据可用性	22
ZFSSA 配置	23
浏览器用户界面 (Browser User Interface, BUI)	24
主窗口	25
一般用法	30
支持的浏览器	35
命令行界面 (Command Line Interface, CLI)	35
登录到 CLI	36
CLI 上下文	36
返回到之前的上下文	39
导航到父上下文	39
上下文和 Tab 补齐功能	40
执行特定于上下文的命令	40
未提交的上下文	40
属性	41
2 状态	45
显示板	46
链路	46
CLI	53
▼ 持续运行 "Dashboard" (显示板)	54
设置	54
简介	54
BUI	55

CLI	57
任务	57
NDMP 状态	57
NDMP 状态 – BUI	58
NDMP 状态 – CLI	60
3 初始配置	61
先决条件	61
使用 BUI 执行初始配置	61
▼ 执行初始配置	62
配置管理端口	62
使用 CLI 执行初始配置	63
4 网络配置	69
网络配置页	69
设备	71
数据链路	71
网络接口	73
网络 IP 多路径 (IPMP)	74
网络性能和可用性	75
网络路由配置	75
使用 BUI 进行网络配置	78
网络配置页	79
网络地址	80
网络路由页	80
使用 CLI 进行网络配置	81
使用 BUI 执行网络配置任务	83
▼ 创建单个端口接口	83
▼ 修改接口	84
▼ 通过拖放创建单个端口接口	84
▼ 创建 LACP 聚合链路接口	84
▼ 创建使用基于探测和链路状态故障检测的 IPMP 组	85
▼ 创建仅使用链路状态故障检测的 IPMP 组	85
▼ 扩展 LACP 聚合	86
▼ 扩展 IPMP 组	86
▼ 创建 InfiniBand 分区数据链路和接口	86
▼ 为群集控制器创建不带 VLAN ID 的 VNIC	87
▼ 为群集控制器创建带有相同 VLAN ID 的 VNIC	88
▼ 添加静态路由	89

▼ 删除静态路由	89
使用 CLI 执行网络配置任务	89
▼ 添加静态路由	89
▼ 删除静态路由	89
▼ 将多宿主属性更改为 strict	90
5 存储配置	91
存储配置文件	92
存储配置规则和准则	93
存储验证	93
SAS-2 系统上的存储分配	94
数据配置文件配置	94
导入现有存储池	96
添加其他存储	96
取消配置存储	96
存储池清理	97
使用 BUI 配置存储	97
▼ 配置存储池	97
▼ 将高速缓存设备添加到现有池中	97
使用 CLI 配置存储	98
▼ 将高速缓存设备添加到现有池中	98
6 存储区域网络配置	101
SAN 目标和启动器	101
SAN 目标和启动器组	101
使用 BUI 配置 SAN	102
使用 CLI 配置 SAN	103
SAN 术语	103
SAN 光纤通道	105
FC 端口目标配置	105
FC 启动器配置	106
性能注意事项	106
FC 故障排除	107
使用 BUI 配置 FC	108
使用 CLI 配置 FC	112
iSCSI	115
目标配置	115
启动器配置	116
规划客户机配置	117

iSCSI 故障排除	117
监测 iSCSI 性能	118
使用 BUI 配置 iSCSI	118
使用 CLI 配置 iSCSI	121
SRP	122
SRP 目标配置	122
启动器配置	123
监测 SRP 性能	123
使用 BUI 配置 SRP 目标	124
使用 CLI 配置 SRP 目标	124
7 用户配置	127
用户角色	127
用户授权	128
管理用户属性	129
用户属性	129
角色属性	129
用户 BUI 页面	130
使用 BUI 配置用户	130
▼ 添加管理员	130
▼ 添加角色	131
▼ 为角色添加授权	131
▼ 从角色中删除授权	131
▼ 添加只能查看显示板的用户	131
使用 CLI 配置用户	132
CLI 用户配置示例	132
▼ 添加管理员	133
▼ 添加角色	134
▼ 为角色添加授权	134
▼ 从角色中删除授权	135
8 设置 ZFSSA 首选项	137
首选项属性	137
使用 CLI 设置首选项	137
使用 CLI 设置 SSH 公钥	138
9 警报配置	141
警报类别	141
支持的警报操作	142

阈值警报	143
使用 BUI 配置警报	144
使用 CLI 配置警报	145
10 群集配置	149
群集的功能与优势	149
群集的缺点	150
群集术语	150
了解群集	151
群集互连 I/O	151
了解群集资源管理	154
群集的接管和故障恢复	156
群集环境中的配置更改	158
存储的群集注意事项	158
网络的群集注意事项	160
专用本地 IP 接口	161
Infiniband 的群集注意事项	161
群集冗余路径方案	162
防止“记忆分裂”情况	163
估计和降低接管影响	165
使用 BUI 进行群集配置	167
▼ 配置群集	167
▼ 取消群集配置	168
使用 CLI 配置群集	169
▼ 关闭群集配置	169
▼ 关闭备用机头	170
▼ 取消群集配置	171
群集节点布线	171
存储机框布线	173
群集配置 BUI 页面	173
11 ZFSSA 服务	175
可用服务	175
数据服务	176
目录服务	177
服务设置	177
远程访问服务	178
安全服务	178
所需的最少端口数	178

使用 BUI 配置服务	179
▼ 查看特定服务屏幕	180
▼ 查看特定服务屏幕	180
▼ 启用服务	180
▼ 禁用服务	180
▼ 定义属性	180
▼ 查看服务日志	181
使用 CLI 配置服务	181
▼ 选择服务	183
▼ 查看服务状态	183
▼ 启用服务	183
▼ 禁用服务	183
▼ 设置属性	184
▼ 查看服务帮助	184
NFS	185
属性	185
Kerberos Realms (Kerberos 领域)	186
服务日志	187
NFS Analytics	187
NFS BUI 和 CLI 属性	188
▼ 通过 NFS 共享文件系统	189
iSCSI 服务	189
iSCSI 服务属性	189
iSCSI 验证服务	190
iSCSI 服务授权	190
iSCSI 服务目标和启动器	190
iSCSI 故障排除	190
SMB 服务	191
SMB 服务属性	191
SMB 共享资源属性	192
NFS/SMB 互操作性	192
SMB DFS 名称空间	193
▼ 示例：操作 DFS 名称空间	195
SMB 自动主目录服务	195
SMB 本地组	197
SMB 本地帐户	198
SMB MMC 集成	198
使用 BUI 配置 SMB	203
FTP 服务	206
FTP 属性	206

FTP 日志	207
使用 BUI 配置 FTP	207
HTTP 服务	208
HTTP 属性	208
HTTP 验证和访问控制	208
HTTP 日志	209
配置 HTTP	209
NDMP 服务	209
NDMP 本地配置与远程配置	210
NDMP 备份格式和类型	210
NDMP 增量备份	213
NDMP 属性	214
NDMP 日志	215
远程复制	216
影子迁移	216
影子迁移属性	216
SFTP 服务	216
SFTP 属性	217
SFTP 端口	217
SFTP 日志	217
配置 SFTP	217
SRP 服务	219
TFTP 服务	219
TFTP 属性	220
配置 TFTP	220
病毒扫描服务	220
病毒扫描属性	221
病毒扫描日志	222
配置病毒扫描	223
NIS 服务	223
NIS 属性	223
NIS 日志	224
配置 NIS	224
LDAP 服务	225
LDAP 属性	225
LDAP 定制映射	226
LDAP 日志	227
配置 LDAP	227
Active Directory	228
Active Directory 属性	228

Active Directory 域和工作组	229
Active Directory LDAP 签名	229
Active Directory Windows Server 2012 支持	230
Active Directory Windows Server 2008 支持	230
使用 BUI 配置 Active Directory	231
使用 CLI 配置 Active Directory	232
身份映射服务	233
身份映射属性	233
身份映射规则	234
身份映射映射	235
身份映射日志	235
身份映射最佳实践	235
身份映射概念	236
身份映射示例	237
配置身份映射	238
DNS 服务	239
DNS 属性	239
配置 DNS	240
DNS 日志	240
Active Directory 和 DNS	240
非 DNS 解析	241
无 DNS 运行	241
动态路由服务	241
RIP 和 RIPng 动态路由协议	241
动态路由日志	242
IPMP 服务	242
IPMP 属性	242
IPMP 日志	242
NTP 服务	243
NTP 属性	243
NTP BUI 时钟	244
NTP 提示	244
使用 BUI 配置 NTP	244
使用 CLI 配置 NTP	245
回拨服务	246
Oracle 单点登录帐户	246
回拨属性	247
注册设备	247
回拨状态	248
回拨状态	248

回拨日志	249
REST	249
RESTful API	249
服务标签	249
服务标签属性	249
SMTP 服务	250
SMTP 属性	250
SMTP 日志	250
SNMP 服务	251
SNMP 属性	251
SNMP MIB	251
Sun FM MIB	252
Sun AK MIB	252
配置 SNMP	253
Syslog 服务	254
Syslog 属性	254
经典版 Syslog : RFC 3164	255
更新的 Syslog : RFC 5424	255
SYSLOG 消息格式	255
接收器配置示例	257
系统标识	258
系统标识属性	259
系统标识日志	259
SSH 服务	259
SSH 属性	259
SSH 日志	260
配置 SSH	260
 12 共享资源、项目和模式	 261
了解共享资源	262
存储池	262
使用共享资源	263
共享资源属性	263
共享资源快照	264
共享资源克隆	264
共享资源空间管理	265
共享资源空间术语	265
了解快照	266
文件系统和项目设置	267

用户和组设置	268
文件系统名称空间	271
嵌套了挂载点的名称空间	271
挂载点的协议访问名称空间	272
Shares (共享资源) > Shares (共享资源)	273
在 BUI 中使用 "Shares" (共享资源) > "Shares" (共享资源)	273
在 CLI 中使用 "Shares" (共享资源) > "Shares" (共享资源)	278
"Shares" (共享资源) > "Shares" (共享资源) > "General" (常规) - BUI 页面	284
Space Usage (空间使用情况)	284
Mountpoint (挂载点)	285
Read only (只读)	285
Update access time on read (读取时更新访问时间)	285
Non-blocking mandatory locking (非阻塞强制性锁定)	285
Data deduplication (重复数据删除)	286
Data compression (数据压缩)	286
Checksum (校验和)	287
Cache Device Usage (高速缓存设备使用)	287
Synchronous Write Bias (同步写入偏向)	288
Database Record Size (数据库记录大小)	288
Additional replication (额外复制)	289
Virus scan (病毒扫描)	289
Prevent Destruction (防止销毁)	290
Restrict Ownership Change (限制所有权更改)	290
定制属性	290
"Shares" (共享资源) > "Shares" (共享资源) > "Protocols" (协议) - BUI 页面	290
共享资源协议	290
共享资源协议 - NFS	291
共享资源 - SMB	295
共享资源 - iSCSI	296
共享资源 - HTTP	296
共享资源 - FTP	297
共享资源 - SFTP	297
Shares (共享资源) > Shares (共享资源) > Access (访问)	297
访问控制	297
共享资源 - 根目录访问权限	297
共享资源 - ACL 行为	299
根目录 ACL	301
共享资源 - 快照	303

共享资源 – 快照属性	303
使用 BUI 列出快照	304
使用 BUI 创建手动快照	305
使用 BUI 调度的快照	307
使用 CLI 创建手动快照	308
项目	312
通过 BUI 使用项目	312
通过 CLI 使用项目	314
项目 – 常规	317
项目协议	319
项目访问	320
项目快照	320
模式	322
定制的共享资源属性	322
在 BUI 中使用模式	322
通过 CLI 使用模式	323
▼ 使用 CLI 配置模式	324
13 复制	325
复制概述	325
了解复制	326
复制术语	326
项目复制目标	327
项目复制操作和数据包	327
项目复制存储池	328
项目级与共享资源级复制	329
配置项目复制	329
创建和编辑目标	330
▼ 在 BUI 中创建和编辑目标	330
▼ 在 CLI 中创建和编辑目标	330
创建和编辑操作	331
▼ 在 BUI 中创建和编辑操作	332
▼ 在 CLI 中创建和编辑操作	333
复制模式：调度或连续	335
复制 – 包括中间快照	335
复制 – 发送和取消更新	335
管理复制数据包	335
管理 BUI 中的复制数据包	336
在 CLI 中管理复制数据包	338

取消复制更新	338
禁用数据包	339
克隆数据包或各个共享资源	339
导出已复制的文件系统	340
断开复制连接	341
反转复制方向	341
销毁复制数据包	343
复制任务	343
反转复制 - 建立复制	343
▼ 反转复制	343
反转复制 - 模拟从灾难中恢复	344
▼ 反转复制	345
反转复制 - 从生产系统恢复复制	346
▼ 反转复制	346
强制复制以使用静态路由	347
▼ 强制复制以使用静态路由	347
克隆接收的复制项目	350
远程复制详细信息	351
授权	351
警报	351
复制审计事件	351
复制和群集	352
快照和数据一致性	352
快照管理	353
复制 iSCSI 配置	354
复制克隆	354
监测复制	354
复制故障	355
复制兼容性	356
从 2009.Q3 及更早版本升级	357
 14 影子迁移	 359
数据迁移	359
传统数据迁移	359
影子迁移	360
影子迁移的行为方式	361
影子源的限制	361
迁移过程中的影子文件系统语义	361
标识和 ACL 迁移	362

影子迁移管理	363
创建影子文件系统	363
管理后台迁移	363
处理迁移错误	363
监视迁移进度	364
取消迁移	364
影子文件系统的快照	365
备份影子文件系统	365
复制影子文件系统	365
影子迁移分析	365
迁移本地文件系统	366
影子迁移任务	366
▼ 测试潜在的影子迁移	366
▼ 从活动 NFS 服务器迁移数据	367
 15 CLI 脚本化	369
自动化访问	369
批处理命令	369
脚本化命令	370
脚本环境	370
与系统交互	370
生成输出	375
处理错误	375
 16 维护工作流	377
使用工作流	377
工作流执行上下文	378
工作流参数	378
约束参数	380
可选参数	380
工作流错误处理	381
工作流输入验证	382
工作流执行审计	383
工作流执行报告	383
版本控制	384
设备版本控制	385
工作流版本控制	385
作为警报操作的工作流	385
警报操作执行上下文	386

审计警报操作	386
使用调度的工作流	387
使用 CLI	387
对调度表进行编码	388
示例：设备类型选择	390
BUI	392
CLI	393
下载工作流	393
查看工作流	394
执行工作流	394
17 集成	397
Oracle Exadata Database Machine 备份	398
手动配置 Sun ZFS Storage Appliance	398
配置网络、池和共享资源	398
配置 Oracle RMAN 和 Oracle 数据库实例	400
后续步骤	401
针对 Sun ZFS Storage Appliance 配置 Oracle Exadata	401
配置 Exadata：为 Sun ZFS Storage Appliance 配置 Oracle Exadata	401
具体实施步骤	402
Oracle SPARC SuperCluster 备份	406
配置 ZFS Storage Appliance 以便备份	406
配置 ZFS Storage Appliance InfiniBand 数据链路	407
配置 Oracle SPARC SuperCluster InfiniBand 交换机以添加 ZFS Storage Appliance	408
为单个 IP 连接配置 ZFS Storage Appliance 网络	410
配置 ZFS Storage Appliance 网络以实现主动-主动配置	411
配置 ZFS Storage Appliance 存储池	413
配置 ZFS Storage Appliance 共享资源	413
配置 ZFS Storage Appliance DTrace Analytics	414
配置客户机 NFS 挂载	415
调整 Solaris 11 网络和内核	415
配置 Oracle Direct NFS (dNFS)	416
为 Oracle RMAN 备份和恢复操作调整 Oracle 数据库实例	417
为 Oracle RMAN 操作创建专用服务	419
配置 Oracle RMAN	419
后续步骤	424
为 ZFS Storage Appliance 备份配置 Oracle SPARC SuperCluster	424
配置 SSC：为 ZFS Storage Appliance 备份配置 Oracle SPARC SuperCluster	424

具体实施步骤	425
Oracle Intelligent Storage Protocol	428
设置最佳的文件记录大小	428
为每个请求使用 ZFS Latency (延迟) 或 Throughput (吞吐量) 写入模 式	428
适用于 Oracle Solaris Cluster 的 Sun ZFS Storage Appliance 网络文件系统插 件	429
用于 Oracle Solaris Cluster Geographic Edition 的 Sun ZFS Storage Appliance 插件	429
适用于 Oracle Enterprise Manager Grid Control 的 Sun ZFS Storage 管理插 件	429
适用于 Oracle Enterprise Manager Grid Controller 的 Oracle Grid Controller Sun ZFS Storage 管理插件	430
适用于 Sun ZFS Storage Appliance 的 Oracle 虚拟机存储连接插件	431
适用于卷影副本服务软件的 Sun ZFS Storage Appliance 提供程序	431
对 Symantec DMP/Storage Foundation 的 FC 支持	432
针对以下 OS 版本提供对 Symantec Storage Foundation 5.1RP2 及更高 版本的 FC 支持	432
适用于 VMware Site Recovery Manager 的 Sun ZFS Storage 7000 Storage Replication Adapter	433
索引	435

使用本文档

- 概述 – 介绍如何管理 Oracle ZFS Storage Appliance
- 目标读者 – 技术人员、系统管理员和授权服务提供商
- 必备知识 – 使用 Oracle ZFS Storage Appliance 的经验

产品文档库

要获取 Oracle ZFS Storage Appliance 文档库，请访问产品文档库，网址为：<http://www.oracle.com/goto/ZFSStorage/docs>。

要获取相关文档（包括白皮书），请访问 <http://www.oracle.com/technetwork/server-storage/sun-unified-storage/overview/index.html> 并单击 "Documentation"（文档）选项卡。有关本产品的最新信息和已知问题，请访问 My Oracle Support，网址为：<http://support.oracle.com>。

获得 Oracle 支持

Oracle 客户可通过 My Oracle Support 获得电子支持。有关信息，请访问 <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info>；如果您听力受损，请访问 <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs>。

反馈

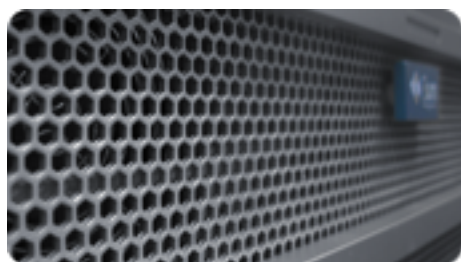
可以在 <http://www.oracle.com/goto/docfeedback> 上提供有关此文档的反馈。

◆◆◆ 第 1 章

Oracle ZFS Storage Appliance 概述

Oracle ZFS Storage Appliance (ZFSSA) 系列产品可通过网络为客户机提供高效率的文件和块数据服务，并且还可为系统上存储的数据提供一组丰富的数据服务。

ZFSSA 主要功能



Oracle ZFS Storage 系统采用了一些技术，以便实现最佳的存储性价比和无与伦比的生产环境工作负荷监测能力，这些新技术包括：

- [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Analytics”，一种用于实时、动态监测系统行为并以图形方式呈现数据的系统
- ZFS 混合存储池，由可加快读写速度的可选闪存设备、低能耗、大容量磁盘以及 DRAM 内存组成，所有这些部件都作为一个数据分层结构来进行透明管理
- 对各种 [《Oracle ZFS Storage Appliance 客户服务手册》](#) 中的“硬件视图”的支持
- 对各种 [《Oracle ZFS Storage Appliance 客户服务手册》](#) 中的“硬件视图”的支持

支持的协议

ZFSSA 支持各种行业标准客户机协议，其中包括：

- [“SMB” \[191\]](#)

- [“NFS” \[185\]](#)
- [“HTTP 和 HTTPS” \[208\]](#)
- [“WebDAV” \[208\]](#)
- [“iSCSI” \[189\]](#)
- [“SAN 光纤通道” \[105\]](#)
- [“SRP” \[122\]](#)
- [iSER 目标配置 \[118\]](#)
- [“FTP” \[206\]](#)
- [“SFTP” \[216\]](#)

ZFSSA 数据服务

要管理使用以上协议导出的数据，可使用内置的一组高级数据服务来配置 ZFSSA，这些数据服务包括：

许可声明：可以免费试用远程复制和克隆特性，但若要在生产环境中使用，必须为各特性购买单独的许可证。在试用期后，必须为这些特性获取许可证或停用它们。*Oracle* 保留随时审计许可合规性的权利。有关详细信息，请参阅 *“Oracle Software License Agreement (“SLA”) and Entitlement for Hardware Systems with Integrated Software Options”*。

- RAID-Z (RAID-5 和 RAID-6)、镜像以及条带化 [第 5 章 存储配置](#)
- 无限制只读和读写 [“共享资源 – 快照” \[303\]](#) 以及快照调度
- [“重复数据删除” \[284\]](#)
- 内置 [“数据压缩” \[284\]](#)
- 用于灾难恢复数据的 [第 13 章 复制](#)
- 具备高可用性的主动-主动 [第 10 章 群集配置](#)
- [“iSCSI” \[189\]](#) “LUN”的瘦置备
- [“病毒扫描和隔离” \[220\]](#)
- [“NDMP 备份和恢复” \[209\]](#)

数据可用性

为了最大程度地提高生产环境中数据的可用性，ZFSSA 提供了一个完整的端到端数据完整性体系结构，其中包括堆栈每一级别的冗余。主要功能包括：

- 对所有系统硬件（CPU、DRAM、I/O 卡、磁盘、风扇、电源）故障进行预测性自我恢复和诊断

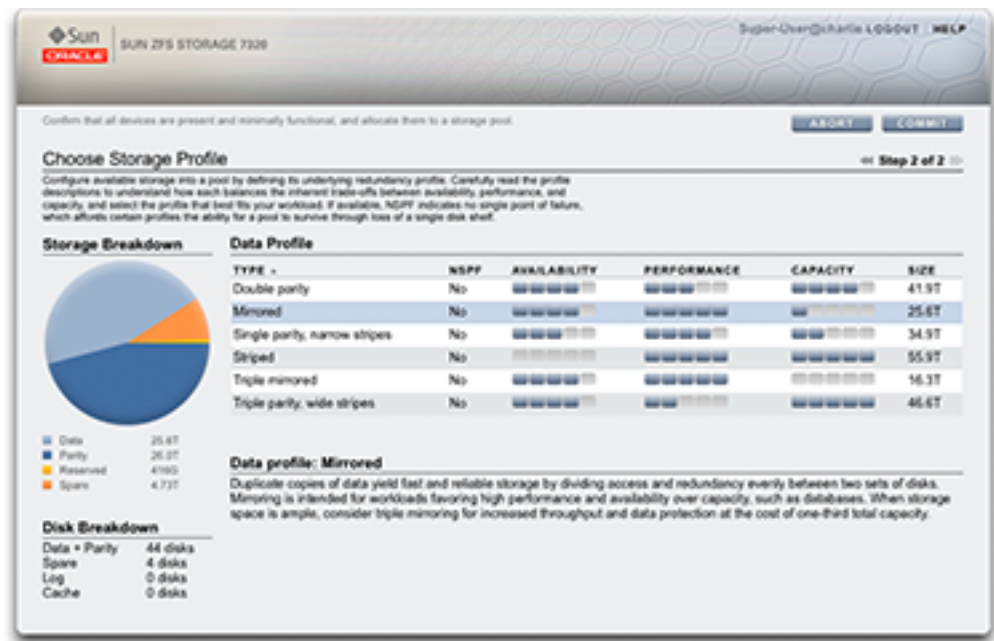
- 为所有数据和元数据提供 ZFS 端到端数据校验和，用于保护整个堆栈的数据
- 为各种磁盘机框配置 RAID-6（双重和三重奇偶校验）和可选的 RAID-6
- 具备高可用性的主动-主动[第 10 章 群集配置](#)
- 用于抵御网络故障的[第 4 章 网络配置](#)
- 在控制器和磁盘机框间提供 I/O 多路径
- 为所有系统[第 11 章 ZFSSA 服务](#)提供了集成的软件重新启动功能
- 为所有软件和硬件问题提供了遥测“回拨”[\[246\]](#)功能
- 可对每个系统进行快速远程管理 (Lights Out Management)，以便进行远程电源控制和控制台访问

ZFSSA 配置

要配置 ZFSSA，请遵循以下各节：

- [第 3 章 初始配置](#) – 初始配置
- [第 4 章 网络配置](#) – 网络
- [第 11 章 ZFSSA 服务](#) – 数据服务
- [第 6 章 存储区域网络配置](#) – 存储区域网络配置
- [第 10 章 群集配置](#) – 群集
- [第 7 章 用户配置](#) – 用户帐户和访问控制
- [第 7 章 用户配置](#) – 用户首选项
- [第 9 章 警报配置](#) – 定制警报
- [第 5 章 存储配置](#) – 重新配置存储设备
- [第 12 章 共享资源、项目和模式](#)

浏览器用户界面 (Browser User Interface, BUI)



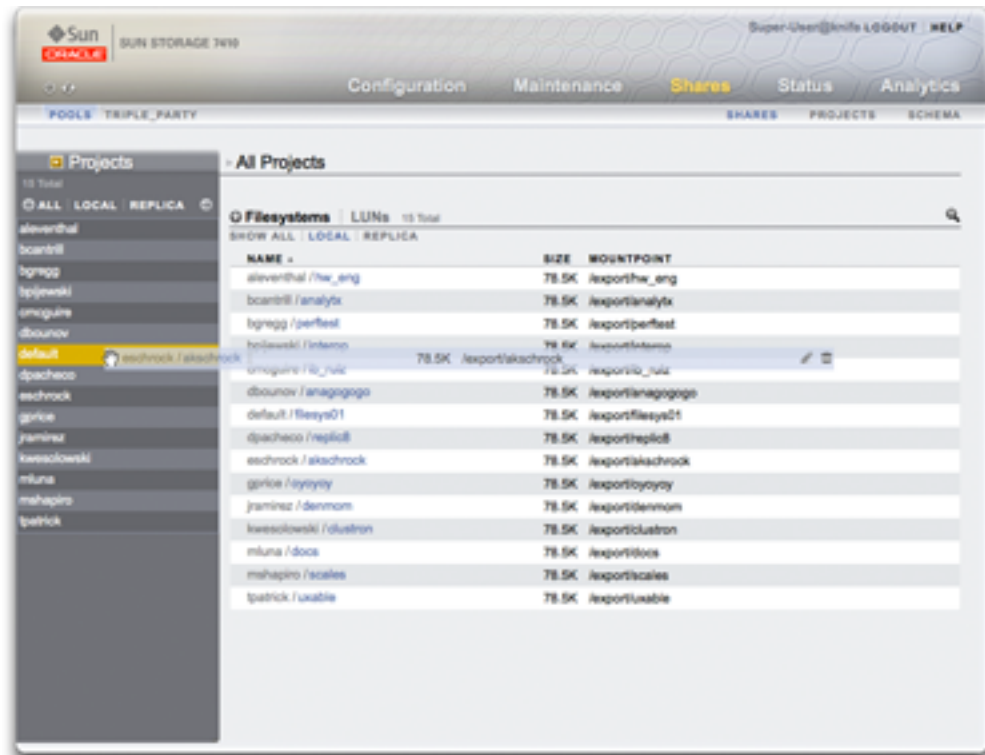
ZFSSA 浏览器用户界面 (Browser User Interface, BUI) 是指用于管理设备的图形工具。BUI 为管理任务、概念描述以及性能数据分析提供了一个直观的环境。BUI 提供了一个用于直观显示系统行为和识别设备性能问题的简洁界面。

使用您在初始配置期间为 NET-0 端口分配的 IP 地址或主机名将您的浏览器定向到该系统，如下所示：<https://ipaddress:215> 或 <https://hostname:215>。此时将显示登录屏幕。

BUI 的右上方提供了与上下文相关的联机帮助链接。对于 BUI 中的每个顶级和二级屏幕，单击 "Help"（帮助）按钮将显示相关的帮助页面。

- “主窗口” [25] – BUI 元素和设计的概述
- “一般用法” [30] – 图标参考
- “支持的浏览器” [35] – 支持的浏览器

主窗口



通过使用 "Projects" (项目) 侧面板将文件系统移到另一个项目中来更改此文件系统的属性。

标头

标头包含用于导航、通知以及主要功能的多种界面元素。左侧从上到下依次为 Sun/Oracle 徽标、硬件型号标签以及硬件关机/重新启动按钮。右侧从上到下依次为：登录标识、注销、帮助、主导航和次级导航。



警报

系统警报在触发时会显示在标头中。如果连续触发多个警报，请参阅[“Dashboard”（显示板）](#) [46] 屏幕上提供的最新警报列表或《Oracle ZFS Storage Appliance 客户服务手册》中的[“日志”](#)屏幕上提供的完整日志。

导航

使用主导航链接在 BUI 的[第 4 章 网络配置](#)、《Oracle ZFS Storage Appliance 客户服务手册》中的[“维护”](#)、[第 12 章 共享资源、项目和模式](#)、[第 2 章 状态](#)和《Oracle ZFS Storage Appliance Analytics 指南》中的[“Analytics”](#)区域之间查看。

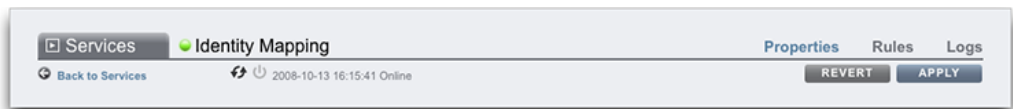
使用次级导航链接可访问每个区域中的特性和功能。

会话注释

如果提供了会话注释，则它会显示在登录 ID 和注销控件的下方。要在不注销的情况下更改后续管理操作的会话注释，请单击文本链接。有关会话注释的详细信息，请参见[第 7 章 用户配置](#)。

标题栏

标题栏显示在标头的下方，提供本地导航以及视当前视图而定的功能。

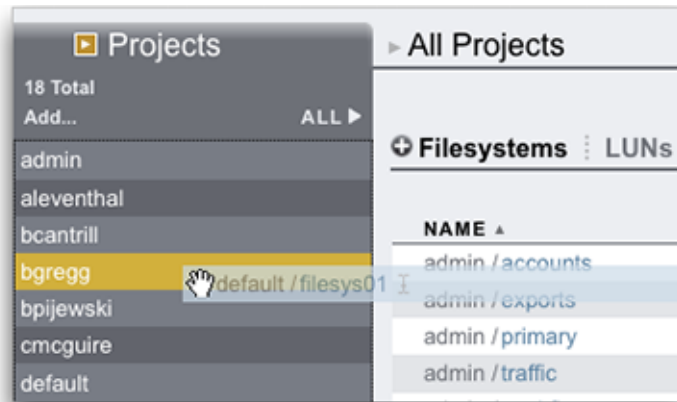


例如，通过身份映射服务标题栏，可以执行以下操作：

- 通过侧面板在完整服务列表中导航
- 通过控件启用或禁用身份映射服务
- 查看身份映射运行时间
- 导航到身份映射服务的 "Properties"（属性）、"Rules"（规则）和 "Logs"（日志）屏幕
- 通过 "Apply"（应用）按钮应用在当前屏幕上所做的配置更改
- 通过 "Revert"（恢复）按钮恢复在当前屏幕上所应用的配置更改

侧面板和菜单标题

要在 "Service"（服务）和 "Project"（项目）视图之间快速导航，请通过单击标题或显示  箭头打开和关闭侧面板。



主窗口侧面板和菜单标题

添加项目

要添加项目，请单击侧边栏中的 "Add..."（添加...）链接。

移动共享资源

要在项目之间移动共享资源，请单击移动  图标，并将某个文件系统共享资源拖到侧面板的适当项目中。

注意将共享资源拖到另一个项目会更改其属性（如果这些属性设置为从其父项目继承）。

对象名称

要更改某个共享资源的名称，请单击该共享资源所在突出显示的表格行中的重命名  图标。

非标准 BUI 控件初识

大部分 BUI 控件都使用标准 Web 表单输入，但还有一些值得注意的重要例外情况：

表 1-1 重要 Web 表单例外情况

BUI 控件摘要	
修改属性	单击编辑  图标并完成对话框
添加一个列表项目或属性条目	单击添加  图标
删除一个列表项目或属性条目	单击删除  图标
保存更改	单击 "Apply" (应用) 按钮
撤消保存的更改	单击 "Revert" (恢复) 按钮
从列表中删除一个项目	单击垃圾箱  图标 (将鼠标悬停在项目行上可以看到该图标)
在列表中搜索某个项目	单击列表右上方的搜索  图标
按列表标题排序	单击粗体的子标题可对列表进行重新排序
移动或拖动某个项目	单击移动  图标
重命名项	单击重命名  图标
查看有关系统的详细信息	Oracle 徽标，或者单击型号标签，这样可转到该型号的 oracle.com 网页
自动打开侧面板	将某个项目拖到侧面板中

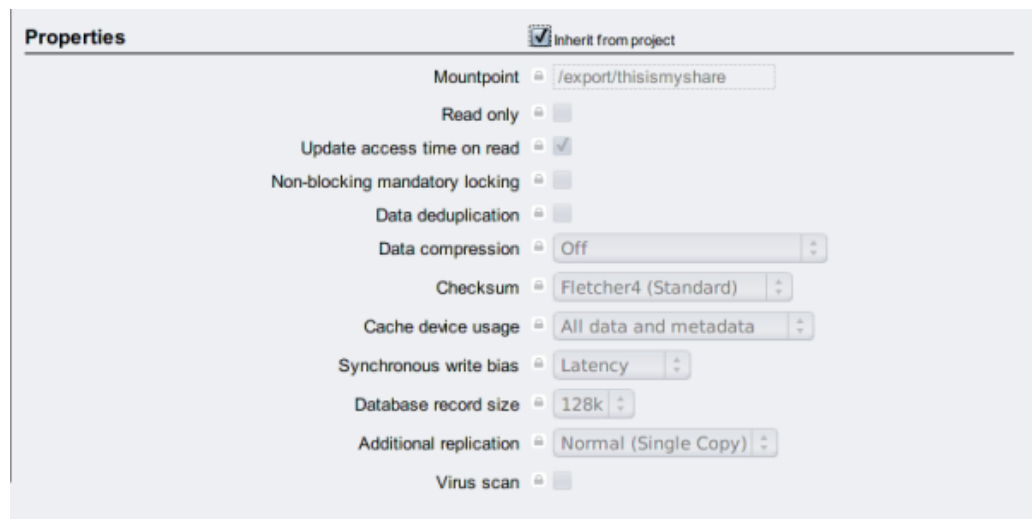
权限

设置权限时，RWX 方块处于可单击状态。单击访问组标签 ("User" (用户)、"Group" (组) 和 "Other" (其他)) 可启用和禁用该标签的所有权限。



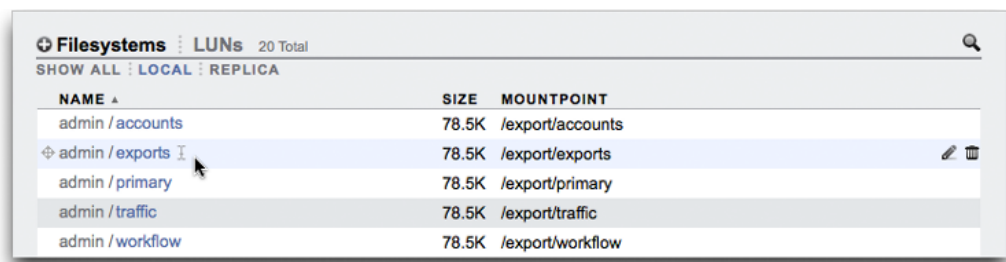
编辑共享资源属性

要编辑共享资源属性，请取消选中 "Inherit from project"（从项目继承）复选框。



查看列表项控件

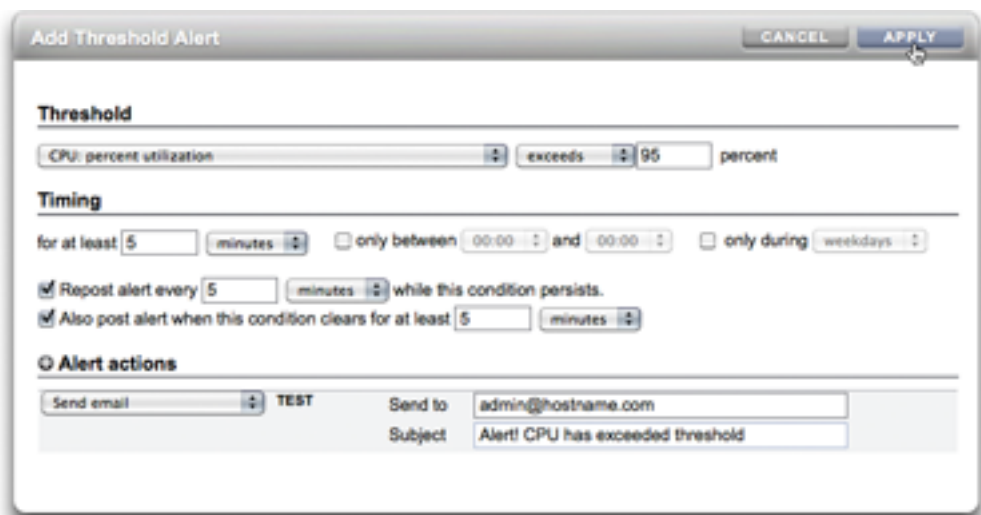
要查看列表中某个项目的控件，请将鼠标悬停在该行上。



NAME	SIZE	MOUNTPOINT
admin / accounts	78.5K	/export/accounts
admin / exports	78.5K	/export/exports
admin / primary	78.5K	/export/primary
admin / traffic	78.5K	/export/traffic
admin / workflow	78.5K	/export/workflow

模态对话框

所有模态对话框在顶部都有标识以及提交或取消当前操作的标题和按钮，内容在下方显示。虽然模态内容区域与主内容区域采用相同的界面约定，但不同的是，必须使用标题栏中的按钮关闭它们后，才能执行其他操作。



Add Threshold Alert [CANCEL] [APPLY]

Threshold

CPU: percent utilization exceeds 95 percent

Timing

for at least 5 minutes ☐ only between 00:00 and 00:00 ☐ only during weekdays

☒ Repost alert every 5 minutes while this condition persists.

☒ Also post alert when this condition clears for at least 5 minutes

Alert actions

Send email TEST Send to admin@hostname.com Subject Alert! CPU has exceeded threshold

一般用法

图标指示系统状态并提供对功能的访问，但是在大多数情况下充当单击时执行操作的按钮。将鼠标悬停在界面图标上非常有用，这样可以查看工具提示。下表提供用户界面约定的要点。

状态

状态指示灯是系统运行状况和服务状态的基本指示器：

表 1-2 状态指示灯

图标	说明	图标	说明
	开		警告
	关		已禁用

基本用法

以下图标显示在用户界面上，并涵盖了大多数基本功能：

表 1-3 BUI 图标

图标*		说明	图标*		说明
--		重命名 (编辑文本)	--		服务器
--		移动	--		克隆
		编辑	--		回滚
		销毁	--		设备电源
		添加	--		应用
		删除	--		恢复
		取消/关闭	--		信息
--		错误	--		对列表列进行排序 (降序)
--		警报	--		对列表列进行排序 (升序)
		打开/关闭开关			第一页
		重新启动			上一页

图标*		说明	图标*		说明
--		定位			下一页
		禁用/脱机			最后一页
		锁定	--		搜索
--		等待自旋动画			菜单
--		调转方向			面板

* 已禁用图标显示在左侧。

联网

以下图标指示网络设备的状态和网络数据链路的类型：

表 1-4 网络图标

图标	说明	图标	说明
	活动网络设备		活动 Infiniband 端口
	非活动网络设备		非活动 Infiniband 端口
	网络数据链路		网络数据链路 (IB 分区)
	网络数据链路 VLAN		
	网络数据链路聚合		
	网络数据链路聚合 VLAN		

显示板阈值

以下图标指示与“[设置](#)” [54]中用户可配置的阈值集有关的监视统计信息的当前状态。

表 1-5 显示板图标

图标	说明	图标	说明
	晴天		飓风
	局部多云		2 级飓风
	多云		3 级飓风
	雨天		4 级飓风
	暴风雨		5 级飓风

Analytics

此组图标在工具栏中用于处理 Analytics 工作表中的信息显示。

表 1-6 Analytics 工具栏图标

图标	说明	图标	说明
	后退		显示最小值
	前进		显示最大值
	前进到现在		显示折线图
	暂停		显示山形图
	缩小		剔除异常值
	放大		将工作表同步到此统计信息
	显示一分钟		不同步工作表统计信息
	显示一小时		深入分析

图标	说明	图标	说明
	显示一天		导出统计数据（下载到客户机）
	显示一周		保存统计数据
	显示一个月		归档数据集
			使用支持包发送工作表

身份映射

以下图标指示在 Windows 和 Unix 之间映射用户和组时应用的角色类型。

表 1-7 身份映射图标

图标*	说明	图标*	说明
	允许 Windows 到 Unix 映射		允许 Unix 到 Windows 映射
	拒绝 Windows 到 Unix 映射		拒绝 Unix 到 Windows 映射
	允许双向映射		

* 已禁用图标显示在左侧。

其他图标

以下图标用于区分不同的对象类型并提供次要信息。

表 1-8 其他图标

图标	说明	图标	说明
	允许		SAS
	拒绝		SAS 端口
	存储池		

支持的浏览器

本节阐述了 BUI 浏览器支持。为了获得最佳效果，请使用第 1 层浏览器。

第 1 层

BUI 软件设计为在以下第 1 层浏览器上可以发挥其全面、完善的功能：

- Firefox 3.x 或更高版本
- Internet Explorer 7 或更高版本
- Safari 3.1 或更高版本
- Google Chrome (稳定版)
- WebKit 525.13 或更高版本

第 2 层

BUI 元素在第 2 层浏览器中可能有些美中不足，某些功能可能不可用，但是所有必要的功能都可以正常工作。如果要使用以下第 2 层浏览器之一，则在登录时会显示一条警告消息：

- Firefox 2.x
- Solaris 10 上的 Mozilla 1.7
- Opera 9

不受支持的浏览器

Internet Explorer 6 以及更早版本都不受支持，已知这些版本存在问题，无法完成登录。

命令行界面 (Command Line Interface, CLI)

CLI 的设计目的在于实现与 BUI 同样的功能，同时也为执行重复性任务提供一个强大的脚本化环境。命令行是用于重复性管理任务的高效而强大的工具。设备通过《[Oracle ZFS Storage Appliance 安装指南](#)》中的“控制台”或“SSH” [259] 提供 CLI。在某些情况下与系统交互的首选方式是通过 CLI，如下所述：

- 网络不可用 – 如果网络不可用，则无法进行基于浏览器的管理；管理的唯一途径就是通过《[Oracle ZFS Storage Appliance 安装指南](#)》中的“控制台”，该控制台只能提供基于文本的界面

- 方便 – 启动浏览器可能非常耗时，特别是在仅需要检查系统的某个特定方面或执行快速配置更改时
- 精确 – 在某些情况下，从本质上讲，浏览器提供的信息与其说是定量信息，不如说是定性信息，而您需要更精确的答案
- 自动 – 基于浏览器的交互无法轻松地自动执行；如果您具有重复性或严格定义的任务，可以编写任务脚本
- Tab 补齐功能的应用非常广泛：如果您不确定在任何给定上下文中应该键入什么内容，则按下 Tab 键可以向您提供可能的选项。在本文档中，按 Tab 键的操作将以粗斜体的 "tab" 一词来表示。
- 可以随时获得帮助：help 命令可提供与上下文相关的帮助。有关特定主题的帮助可以通过将主题指定为 help 的参数来获得，例如 "help commands"。通过使用 Tab 补齐 help 命令或键入 help topics，可显示提供的主题。
-

在 CLI 中导航时，需要注意以下两个原则：

- Tab 补齐功能的应用非常广泛 – 如果您不确定在任何给定上下文中应该键入什么内容，则按下 Tab 键可以向您提供可能的选项。在本文档中，按 Tab 键的操作将以粗斜体的 "tab" 一词来表示。
- 可以随时获得帮助 – help 命令可提供特定于上下文的帮助。有关特定主题的帮助可以通过将主题指定为 help 的参数来获得，例如 help commands。通过使用 Tab 补齐 help 命令或键入 help topics，可显示提供的主题。

这两个原则可以结合起来，如下所示：

```
dory:> help tab
builtins    commands  general    help        properties  script
```

登录到 CLI

要使用 CLI 远程登录，请使用 ssh 客户机。如果没有[第 7 章 用户配置](#)来管理设备，则需要以 root 身份登录。登录时，CLI 会显示一个提示符，其结构为主机名后跟一个冒号，再跟一个大于号：

```
% ssh root@dory
Password:
Last login: Mon Oct 13 15:43:05 2009 from kiowa.sf.fishpo
dory:>
```

CLI 上下文

CLI 中的一个核心原则就是执行命令时所在的上下文。上下文指定可管理系统的哪些元素，以及可使用哪些命令。上下文采用树结构，在该结构中，上下文本身可能包含嵌套的上下文，并且该结构通常是 BUI 中视图结构的镜像。

根上下文

登录时的初始上下文是根上下文，充当所有上下文的父代或祖先。要导航到某个上下文，请将该上下文的名称作为命令执行。例如，在浏览器[第 4 章 网络配置](#)视图中提供的功能在 CLI 的 configuration 上下文中提供。从根上下文中，可以通过直接键入某个上下文来访问该上下文：

```
dory:> configuration
dory:configuration>
```

请注意，提示符将发生更改，在提示符中的冒号和大于号之间提供上下文，以反映上下文。

子上下文

show 命令显示子上下文。例如，从 configuration 上下文中：

```
dory:configuration> show
Children:
    net => Configure networking
    services => Configure services
    version => Display system version
    users => Configure administrative users
    roles => Configure administrative roles
    preferences => Configure user preferences
    alerts => Configure alerts
    storage => Configure Storage
```

这些子上下文对应于在浏览器中的[第 6 章 存储区域网络配置](#)视图下提供的视图，包括[第 4 章 网络配置](#)、[第 11 章 ZFSSA 服务](#)和[第 7 章 用户配置](#)、“首选项”[第 8 章 设置 ZFSSA 首选项](#)等。要选择这些子上下文之一，请键入其名称：

```
dory:configuration> preferences
dory:configuration preferences>
```

可通过指定使用空格分隔的中间上下文直接从祖先导航到某个后代上下文。例如，要直接从根上下文导航到 configuration preferences，直接键入即可：

```
dory:> configuration preferences
dory:configuration preferences>
```

动态子上下文

有些子上下文是动态的，因为它们不是对应于浏览器中的固定视图，而是对应于用户或系统创建的动态实体。要导航到这些上下文，请使用 select 命令，后跟动态上下文的名称。给定上下文中包含的动态上下文的名称可以使用 list 命令显示。例如，users 上下文是一个静态上下文，但是每个用户都是独立的动态上下文。

```
dory:> configuration users
```

```
dory:configuration users> list
NAME                USERNAME            UID      TYPE
John Doe            bmc                 12345    Dir
Super-User          root                0        Loc
```

要选择名为 bmc 的用户，请发出 select bmc 命令：

```
dory:configuration users> select bmc
dory:configuration users bmc>
```

另外，在某些上下文中，可以使用 select 和 destroy 根据实体的属性选择实体。例如，可以通过发出以下命令在 maintenance logs system 上下文中选择 reboot 模块发布的日志条目：

```
dory:maintenance logs system> select module=reboot
dory:maintenance logs system entry-034> show
Properties:
  timestamp = 2010-8-14 06:24:41
  module = reboot
  priority = crit
  text = initiated by root on /dev/console syslogd: going down on signal 15
```

和其他命令一样，select 可以附加到更改上下文的命令中。例如，要从根上下文中选择名为 bmc 的用户：

```
dory:> configuration users select bmc
dory:configuration users bmc>
```

上一个上下文

使用 last 命令导航到上一个选定或创建的上下文。当前仅在复制操作上下文中实施此命令。

下例创建了一个复制操作，然后使用 last 和 get id 命令检索此复制操作 ID。然后选择不同的操作，并使用 last 和 get id 命令检索最后一次访问的复制操作的 ID。

```
dory:shares p1/share replication> list
      TARGET      STATUS      NEXT
action-000 oakmeal      idle      Sync now
action-001 dory         idle      Sync now
dory:shares p1/share replication> create
dory:shares p1/share action (uncommitted)> set target=dory
      target = dory (uncommitted)
dory:shares p1/share action (uncommitted)> set pool=p0
      pool = p0 (uncommitted)
dory:shares p1/share action (uncommitted)> commit
dory:shares p1/share replication> last
dory:shares p1/share action-002> get id
      id = 7034367a-d4d8-e26f-fa93-c3b454e3b595
dory:shares p1/share action-002> done
dory:shares p1/share replication> select action-000
dory:shares p1/share action-000> get id
      id = 9895d9f4-7b23-ebel-faf2-d85a581e3dff
```

```
dory:shares p1/share action-000> done
dory:shares p1/share replication> last get id
id = 9895d9f4-7b23-ebel-faf2-d85a581e3dff
dory:shares p1/share replication>
```

返回到之前的上下文

要返回到之前的上下文，请使用 `done` 命令：

```
dory:configuration> done
dory:>
```

请注意，该命令会导致返回到之前的上下文，该上下文不一定是父上下文，如下所示：

```
dory:> configuration users select bmc
dory:configuration users bmc> done
dory:>
```

可以使用 `done` 命令多次，以返回到更早的上下文：

```
dory:> configuration
dory:configuration> users
dory:configuration users> select bmc
dory:configuration users bmc> done
dory:configuration users> done
dory:configuration> done
dory:>
```

导航到父上下文

要导航到父上下文，请使用 `cd` 命令。受经典 UNIX 命令的启发，`cd` 使用 `".."` 参数表示移动到父上下文：

```
dory:> configuration users select bmc
dory:configuration users bmc> cd ..
dory:configuration users>
```

和 UNIX 命令一样，"`cd /`" 移到根上下文：

```
dory:> configuration
dory:configuration> users
dory:configuration users> select bmc
dory:configuration users bmc> cd /
dory:>
```

和其 UNIX 对应项一样，"`cd ../../`" 可用于导航到祖父上下文：

```
dory:> configuration
dory:configuration> users
dory:configuration users> select bmc
dory:configuration users bmc> cd ../../
dory:configuration>
```

上下文和 Tab 补齐功能

上下文名称可以使用 Tab 补齐功能，上下文可以是静态上下文（通过常规命令补齐功能），也可以是动态上下文（通过 `select` 命令的命令补齐功能）。在下面的示例中，只敲击键盘 15 次即从根上下文中选择了名为 `bmc` 的用户，如果不使用 Tab 补齐功能，则需要敲击键盘 31 次：

```
dory:> configtab
dory:> configuration utab
dory:> configuration users setab
dory:> configuration users select tab
bmc  root
dory:> configuration users select btab
dory:> configuration users select bmccenter
dory:configuration users bmc>
```

执行特定于上下文的命令

进入某个上下文后，可执行特定于上下文的命令。例如，要获取当前用户的首选项，请从 `configuration preferences` 上下文中执行 `get` 命令：

```
dory:configuration preferences> get
      locale = C
      login_screen = status/dashboard
      session_timeout = 15
      session_annotation =
      advanced_analytics = false
```

如果更改上下文的命令后还有其他输入内容，则虽然会在目标上下文中执行该命令，但是控制会返回调用上下文。例如，要在不更改上下文的情况下从根上下文中获取首选项，请将 `get` 命令附加到上下文导航命令：

```
dory:> configuration preferences get
      locale = C
      login_screen = status/dashboard
      session_timeout = 15
      session_annotation =
      advanced_analytics = false
```

未提交的上下文

当在系统中创建新的实体时，与新实体关联的上下文通常以未提交状态创建。例如，通过从 `configuration alerts threshold` 上下文中执行 `create` 命令来创建[第 9 章 警报配置](#)：

```
dory:> configuration alerts thresholds create
dory:configuration alerts threshold (uncommitted)>
```


提示符中的 (uncommitted) 表示这是一个未提交的上下文。未提交的实体通过 `commit` 命令提交；从未提交的上下文导航到其他位置的任何尝试都将提示进行确认：

```
dory:configuration alerts threshold (uncommitted)> cd /
Leaving will abort creation of "threshold". Are you sure? (Y/N)
```

当提交未提交的实体时，将验证与新实体关联的属性，如果无法创建实体，则将生成错误。例如，创建新阈值警报需要指定统计信息名称；不设置该名称会导致错误的发生：

```
dory:configuration alerts threshold (uncommitted)> commit
error: missing value for property "statname"
```

要解决此问题，请修正错误，然后重新尝试提交：

```
dory:configuration alerts threshold (uncommitted)> set statname=cpu.utilization
statname = cpu.utilization (uncommitted)
dory:configuration alerts threshold (uncommitted)> commit
error: missing value for property "limit"
dory:configuration alerts threshold (uncommitted)> set limit=90
limit = 90 (uncommitted)
dory:configuration alerts threshold (uncommitted)> commit
dory:configuration alerts thresholds> list
THRESHOLD      LIMIT      TYPE STATNAME
threshold-000    90      normal cpu.utilization
```

属性

CLI 属性

属性是与上下文关联的类型化名称/值对。给定上下文的属性可以通过运行 "help properties" 命令确定。下面的示例检索与用户首选项关联的属性：

```
dory:configuration preferences> help properties
Properties that are valid in this context:

locale          => Locality

login_screen     => Initial login screen

session_timeout  => Session timeout

session_annotation => Current session annotation

advanced_analytics => Make available advanced analytics statistics
```

获取属性

给定上下文的属性可以使用 `get` 命令检索。下面的示例使用 `get` 命令检索用户的首选项：

```
dory:configuration preferences> get
      locale = C
      login_screen = status/dashboard
      session_timeout = 15
      session_annotation =
      advanced_analytics = false
```

获取单个属性值

get 命令将返回作为参数提供给它的任何属性。例如，要获取 login_screen 属性的值，请运行以下命令：

```
dory:configuration preferences> get login_screen
      login_screen = status/dashboard
```

Tab 补齐功能

get 命令将对可用属性的名称使用 Tab 补齐功能。例如，要查看“[SCSI](#)” [189]服务的可用属性列表，请运行以下命令：

```
dory:> configuration services iscsi get tab
<status>      isns_server      radius_secret      target_chap_name
isns_access    radius_access    radius_server      target_chap_secret
```

设置属性

set 命令使用通过等号分隔的属性名称及其值来将属性设置为指定的值。例如，要将 login_screen 属性设置为 "shares"，请运行以下命令：

```
dory:configuration preferences> set login_screen=shares
      login_screen = shares (uncommitted)
```

请注意，对于构成设备状态的属性，设置属性不会更改其值，而是记录设置的值并指示该属性的值未提交。

提交设置的属性值

要强制设置的属性值生效，必须显式提交这些值，以允许一次连续更改多个值。要提交任何未提交的属性值，请使用 commit 命令：

```
dory:configuration preferences> get login_screen
      login_screen = shares (uncommitted)
dory:configuration preferences> commit
dory:configuration preferences> get login_screen
      login_screen = shares
```

如果尝试离开包含未提交属性的上下文，则系统会警告您离开将放弃设置的属性值，并提示您确认自己将离开。例如：

```
dory:configuration preferences> set login_screen=maintenance/hardware
      login_screen = maintenance/hardware (uncommitted)
dory:configuration preferences> done
You have uncommitted changes that will be discarded. Are you sure? (Y/N)
```

通过隐式提交设置属性值

如果在某一个上下文中设置另一个上下文中的属性（即，如果将 `set` 命令附加到更改上下文的命令），则提交为隐式提交，并且提交在控制权返回发起操作的上下文之前执行。例如：

```
dory:> configuration preferences set login_screen=analytics/worksheets
      login_screen = analytics/worksheets
dory:>
```

将属性设置为值列表

有些属性采用值列表。对于这些属性，应使用逗号分隔列表元素。例如，“[NTP](#)” [\[243\]](#)的 `servers` 属性可设置为 NTP 服务器列表：

```
dory:configuration services ntp> set servers=0.pool.ntp.org,1.pool.ntp.org
      servers = 0.pool.ntp.org,1.pool.ntp.org (uncommitted)
dory:configuration services ntp> commit
```

将属性设置为包含特殊字符的值

如果某个属性值包含逗号、等号、引号或空格，则整个值必须使用引号括起来。例如，要将默认项目的 `sharenfs` 共享资源属性设置为只读属性，但是向主机 “kiowa” 提供读取/写入访问权限，请运行以下命令。有关信息，请参见[第 12 章 共享资源、项目和模式](#)。

```
dory:> shares select default
dory:shares default> set sharenfs="ro,rw=kiowa"
      sharenfs = ro,rw=kiowa (uncommitted)
dory:shares default> commit
```

不可变属性

有些属性是不可变的；您可以获取它们的值，但是无法设置它们的值。尝试设置不可变属性会导致错误的发生。例如，尝试设置默认项目的不可变 `space_available` 属性。有关信息，请参见[第 12 章 共享资源、项目和模式](#)。

```
dory:> shares select default
dory:shares default> get space_available
      space_available = 1.15T
dory:shares default> set space_available=100P
error: cannot set immutable property "space_available"
```

某些其他属性仅在特定的情况下才不可变。对于这些属性，set 命令无效。例如，如果名为 bmc 的用户是一个网络用户，则 fullname 属性将不可变：

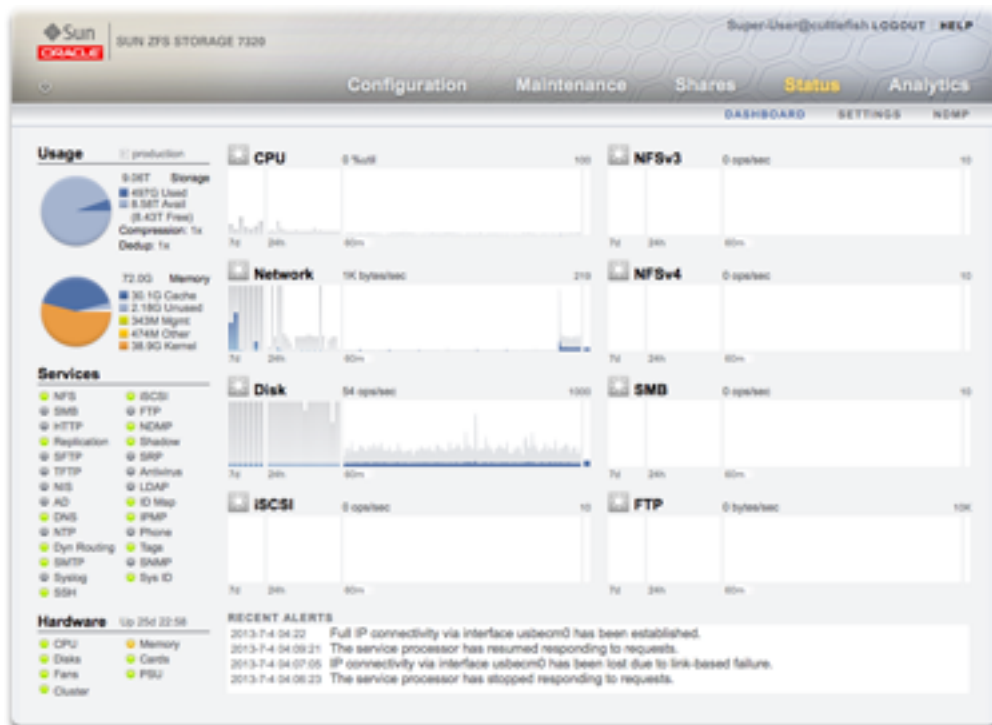
```
dory:> configuration users select bmc set fullname="Rembrandt Q. Einstein"
error: cannot set immutable property "fullname"
```

状态

"Status" (状态) 部分提供设备状态摘要和配置选项。有关设备状态视图和相关服务配置的概念性和程序性信息，请参阅下面的部分：

- [""Status" \(状态\) > "Dashboard" \(显示板\) "](#) [46] 屏幕提供存储、内存、服务、硬件、活动和最新警报的视图。
- [""Status" \(状态\) > "Settings" \(设置\) "](#) [54] 屏幕用于更改 "Dashboard" (显示板) 上显示的图形，并定制 "Dashboard" (显示板) 上每个图形显示的天气图标的关联阈值设置。
- [""Status" \(状态\) > NDMP" \[57\]](#) 屏幕提供配置的所有 NDMP 设备和每个 NDMP 会话近期活动的视图。

显示板



"Dashboard"（显示板）汇总了设备状态

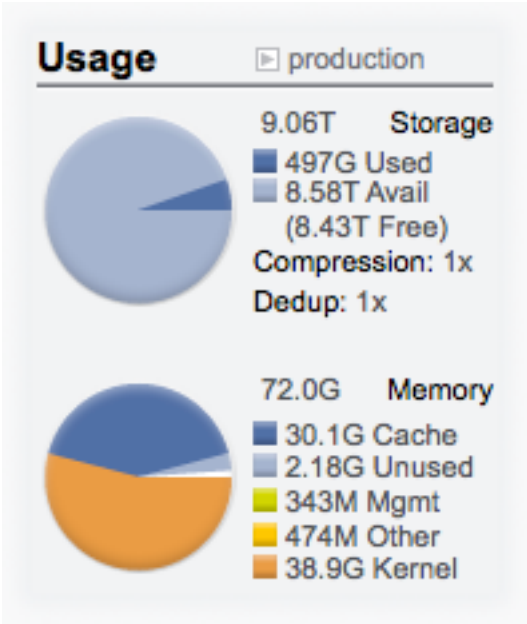
链路

状态显示板提供浏览器用户界面 (browser user interface, BUI) 的所有主屏幕的链接。"Dashboard"（显示板）上有超过 100 个可见项目，它们链接到鼠标悬停时显示的边框或突出显示文本所指示的关联 BUI 屏幕。以下几部分对 "Dashboard"（显示板）的各个区域进行了详细介绍。

使用情况

"Dashboard"（显示板）的 "Usage"（使用情况）区域提供存储池和主内存使用情况的摘要。池名称显示在 "Usage"（使用情况）区域的右上方。如果配置了多个池，请使用下拉列表选择需要显示的池。

图 2-1 状态显示板使用情况



存储

总池容量显示在该区域顶部。"Storage"（存储）饼图详细列明了已用空间和可用空间。要转到池的 "Shares"（共享资源）屏幕，请单击 "Storage"（存储）饼图。

内存

总系统物理内存显示在该区域顶部。左边是按组件显示内存使用情况的饼图。要转到按应用程序名称细分的动态内存使用情况 Analytics 工作表，请单击 "Memory"（内存）饼图。

表 2-1 池使用情况汇总

池使用情况汇总	
Used（已使用）	该池已使用的空间，包括数据和快照。
Avail（可用）	可用物理磁盘空间量。可用于文件数据的空间（在 "Shares"（共享资源）屏幕中报告）将小于这个值，因为文件系统元数据要占用一部分空间。

池使用情况汇总	
Free (空闲)	LUN 容量中的可用空间量减去由项目和共享资源在池中预留的未使用空间。在提前按预留分配磁盘空间和/或创建 LUN 时，提供可用磁盘空间。
Compression (压缩)	该池当前达到的压缩比率。如果禁用了压缩，比率将显示 1x。
Dedup (重复数据删除)	该池当前达到的重复数据删除率。如果禁用了重复数据删除，比率将显示 1x。

表 2-2 主内存使用情况汇总

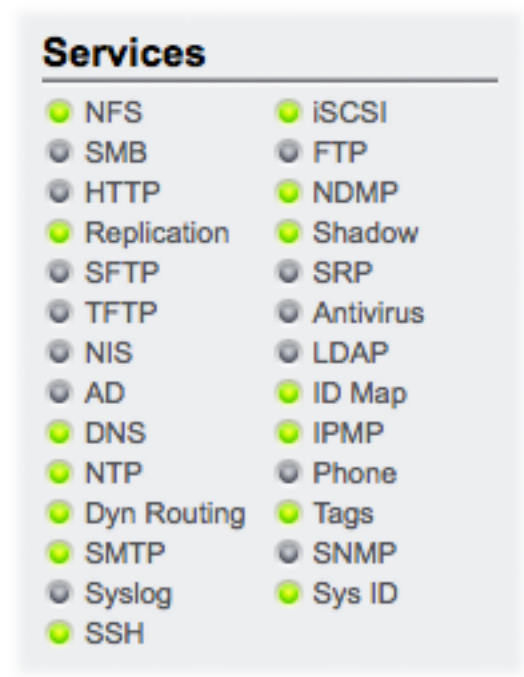
主内存 (RAM) 使用情况汇总	
Cache (高速缓存)	文件系统高速缓存为提高性能而使用的字节数。
Unused (未使用)	当前未使用的字节数。引导后，该值将降低，因为文件系统高速缓存会使用一部分空间。
Mgmt (管理)	设备管理软件使用的字节数。
Other (其他)	其他操作系统软件使用的字节数。
Kernel (内核)	操作系统内核使用的字节数。

请注意，用户需要 analytics/component create+read 授权才能查看内存使用情况。没有该授权，内存详细信息不会显示在 "Dashboard"（显示板）上。

服务

"Dashboard"（显示板）的此区域显示设备上服务的状态，并使用一个指示灯图标显示每项服务的状态。

图 2-2 服务显示板



图标

大多数服务都将显示绿色，指示服务处于联机状态；或者显示灰色，指示服务已禁用。有关所有可能状态和图标颜色的参考信息，请参见[“一般用法” \[30\]](#)一节。

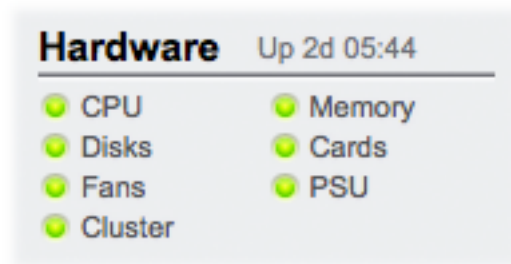
链路

要转到关联的配置屏幕，请单击服务名称。“Properties”（属性）屏幕随即显示，并提供可配置的字段、重新启动、启用和禁用图标，以及该服务关联的“Logs”（日志）屏幕的链接。

硬件

“Dashboard”（显示板）的此区域显示设备上硬件的概述。

图 2-3 硬件显示板



故障

如果存在已知故障，则显示琥珀色故障  图标。

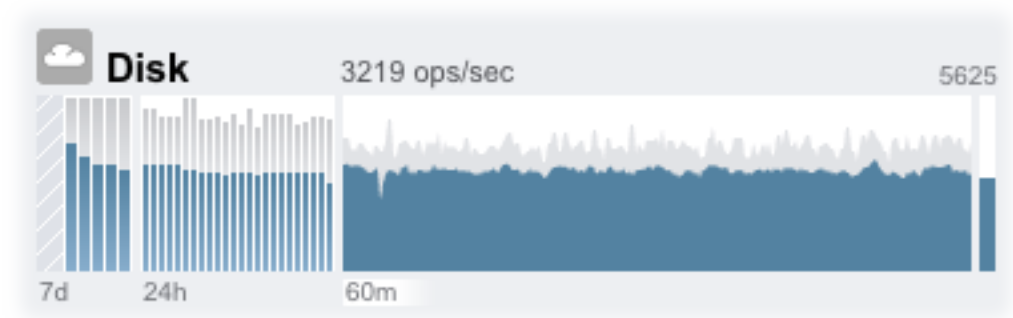
链路

要转到提供硬件状态详细信息的《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“[硬件](#)”屏幕，请单击硬件组件的名称。

活动

"Dashboard"（显示板）的活动区域默认显示八个性能统计信息图形。此部分中的示例显示每秒的磁盘操作数。统计平均值使用蓝色绘制，而最大值则以淡灰色显示。

图 2-4 磁盘活动显示板



要转到活动的《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“Analytics”工作表，请单击要评估的统计信息的四个图形（天、小时、分钟、秒）之一。

要查看每个图形的平均值，请将鼠标悬停在该图形上，平均值将会显示在工具提示中。左上方的天气图标根据阈值提供活动报告，您可以在“[Status Settings](#)”（[状态设置](#)）” [54] 屏幕中针对每种统计信息定制阈值。

图形

表 2-3 统计图形摘要

统计图形摘要	
7 天图 (7d)	条形图，每个条形代表一天。
24 小时图 (24h)	条形图，每个条形代表一小时。
60 分钟图 (60m)	折线图，表示一小时内的活动（还可视为 24 小时图中第一个代表一小时的条形）。
1 秒图	折线图，表示即时活动报告。

平均值

所选绘图的平均值将以数字方式显示在图形的上方。要更改显示的平均值，请选择所需的平均值 – 7d、24h 或 60m。

垂直刻度

所有图形的垂直刻度都显示在右上方，并且所有图形都缩放到此相同的高度。高度根据所选图形进行计算（加上边距）。高度将根据所选图形中的活动重新缩放，但利用率图形除外，该图形具有 100% 的固定高度。

由于高度可以重新缩放，因此空闲活动的 60 分钟看起来可能类似于繁忙活动的 60 分钟。务必先查看图形的高度，然后再尝试解释它们的含义。

有些统计信息可能不是特别好理解 – 您可能会想，对于您环境中的某台特定设备，每秒 1000 个 NFSv3 操作是繁忙还是空闲。这时 24 小时和 7 天绘图就有用处了，它们可以提供临近当前活动的历史数据以供比较。

绘图高度根据所选绘图进行计算。默认情况下，选择 60 分钟绘图。所以，高度是 60 分钟间隔期间的最大活动（加上边距）。要重新缩放所有绘图，以便以过去 7 天期间的最高活动为高度，请选择 "7d"。这可以方便地查看当前活动与过去一天或者一周的比较情况。

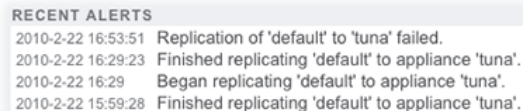
天气

天气图标旨在当存在异常繁忙或空闲的活动时引起您的注意。要转到天气阈值配置页面，请单击天气图标。阈值没有好坏之分，它只是 BUI 为每种活动统计信息提供的级别梯度。通过天气图标所基于的统计信息可以大致了解设备性能，您应该根据自己的工作负荷进行定制，如下所示：

- 不同的环境具有不同的可接受性能（延迟）级别，所以没有统一的阈值。
- "Dashboard"（显示板）上的统计信息基于每秒操作数和每秒字节数，所以应该使用《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“Analytics”工作表来准确了解系统性能。

Recent Alerts（最新警报）

图 2-5 Recent Alerts（最新警报）



RECENT ALERTS	
2010-2-22 16:53:51	Replication of 'default' to 'tuna' failed.
2010-2-22 16:29:23	Finished replicating 'default' to appliance 'tuna'.
2010-2-22 16:29	Began replicating 'default' to appliance 'tuna'.
2010-2-22 15:59:28	Finished replicating 'default' to appliance 'tuna'.

此部分显示最新的四个设备警报。单击该框可转到《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“日志”屏幕以详细查看所有最近的警报。

CLI

"Status" (状态) > "Dashboard" (显示板) 屏幕的文本版本可以通过在 CLI 中键入 status dashboard 获得：

```
cuttlefish:> status dashboard
Storage:
  pool_0:
    Used      497G bytes
    Avail     8.58T bytes
    Free      8.43T bytes
    State      online
    Compression 1x

Memory:
  Cache      30.1G bytes
  Unused     2.18G bytes
  Mgmt       343M bytes
  Other      474M bytes
  Kernel     38.9G bytes

Services:
  ad          disabled      smb          disabled
  dns         online        ftp          disabled
  http        online        identity     online
  idmap       online        ipmp         online
  iscsi       online        ldap         disabled
  ndmp        online        nfs          online
  nis         online        ntp          online
  routing     online        scrk         maintenance
  snmp        online        ssh          online
  tags        online        vscan        online

Hardware:
  CPU         online      Cards        online
  Disks       faulted   Fans         online
  Memory      online    PSU          online

Activity:
  CPU         1 %util    Sunny
  Disk        32 ops/sec Sunny
  iSCSI       0 ops/sec Sunny
  NDMP        0 bytes/sec Sunny
  NFSv3       0 ops/sec Sunny
  NFSv4       0 ops/sec Sunny
  Network     13K bytes/sec Sunny
  SMB         0 ops/sec Sunny

Recent Alerts:
  2013-6-15 07:46: A cluster interconnect link has been restored.
```

前面“BUI” [46]部分中的说明同样适用，但存在以下差异：

- 活动绘图不以文本方式呈现（尽管我们考虑过使用 aalib）。

- 对于存储使用情况部分，CLI 中将列出所有可用池的详细信息，而 BUI 中只有汇总一个池的空间。

还提供了单独的视图，例如 `status activity show`：

```
caji:> status activity show
Activity:
CPU          10 %util          Sunny
Disk         478 ops/sec       Partly Cloudy
iSCSI        0 ops/sec          Sunny
NDMP         0 bytes/sec        Sunny
NFSv3        681 ops/sec          Partly Cloudy
NFSv4        0 ops/sec          Sunny
Network      22.8M bytes/sec       Partly Cloudy
SMB          0 ops/sec          Sunny
caji:>
```

▼ 持续运行 "Dashboard" (显示板)

如果在浏览器中持续（全天候）使 "Dashboard" (显示板) 屏幕保持打开状态，则可能会遇到浏览器内存问题。浏览器的大小会有所增加（内存泄漏），需要先将其关闭，然后再重新打开。浏览器非常善于在浏览不同 Web 站点（以及打开和关闭选项卡）时管理内存。问题是 "Dashboard" (显示板) 屏幕持续运行而未关闭，这样会打开和重新打开活动绘图的图像，从而降低图像呈现的性能。

如果在使用 Firefox 时遇到该问题，请按如下方式禁用内存高速缓存：

1. 打开 `about:config`
2. 筛选“内存”
3. 设置 `browser.cache.memory.enable = false`

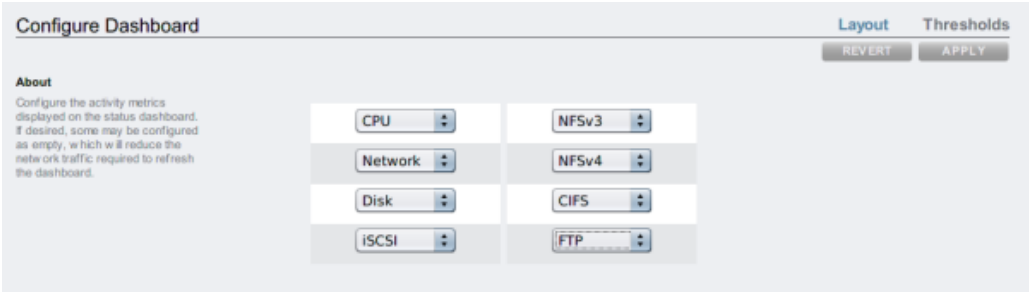
设置

简介

"Status" (状态) > "Settings" (设置) 屏幕允许您定制 ["Status Dashboard" \(状态显示板\)](#) [46]，包括显示的统计信息和通过天气图标指明活动状况的阈值。

BUI

图 2-6 显示板设置



布局

使用 "layout" (布局) 选项卡可选择“[dashboard activity \(显示板活动\)](#)” [46]区域中显示的图形，如下表所定义。

表 2-4 状态布局设置

名称	单位	说明
<empty>	-	此位置将不显示任何图形。
SMB	操作数/秒	SMB 操作的平均数。
CPU	使用率	设备 CPU 处于繁忙状态的平均周期。CPU 周期包括内存等待周期。
Disk (磁盘)	操作数/秒	对物理存储设备进行的平均操作数。
HTTP	操作数/秒	HTTP 操作的平均数。
iSCSI	操作数/秒	iSCSI 操作的平均数。
FC	操作数/秒	光纤通道操作的平均数。
Network (网络)	字节数/秒	所有物理网络接口每秒的平均字节数。
NDMP	字节数/秒	平均 NDMP 网络字节数。
NFSv2	操作数/秒	NFSv2 操作的平均数。

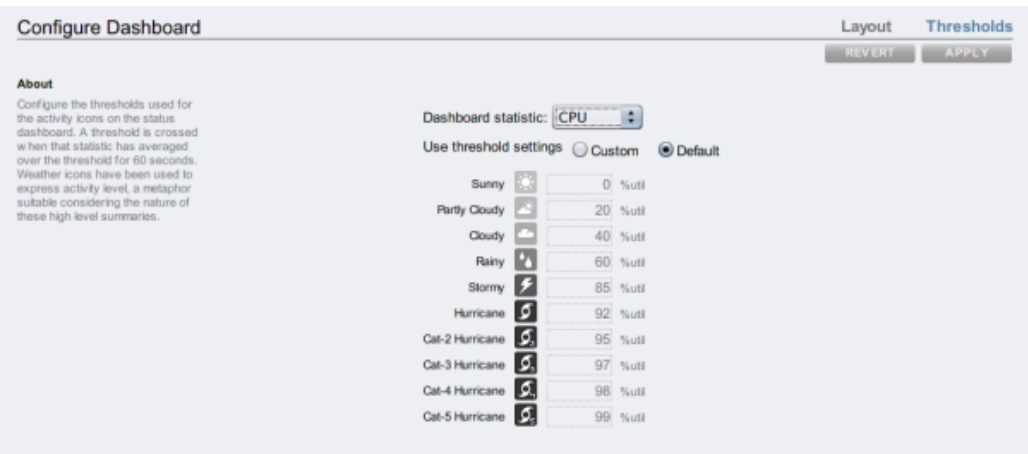
名称	单位	说明
NFSv3	操作数/秒	NFSv3 操作的平均数。
NFSv4	操作数/秒	NFSv4 操作的平均数。
FTP	字节数/秒	FTP 字节的平均数。
SFTP	字节数/秒	SFTP 字节的平均数。

请注意，要减少刷新 "Dashboard" (显示板) 所需的网络流量，请将某些活动图形配置为 "<empty>"。

阈值

使用 "Thresholds" (阈值) 屏幕可配置“显示板活动” [46]的天气图标。提供的默认值以繁重的工作负荷为基础，可能不适合您的环境。

图 2-7 显示板活动设置



“Dashboard” (显示板) [46]上显示的天气图标与当前活动的阈值设置最接近 – 按照 60 秒的平均值进行衡量。例如，如果 CPU 利用率为 41%，则默认情况下，会显示多云天气图标，因为其阈值为 40%（最接近实际活动）。选择 "Custom" (定制) 单选按钮可配置阈值，请务必按照它们在屏幕上的显示顺序进行配置。

CLI

显示板当前无法从 CLI 进行配置。BUI 中保存的设置将应用于在 CLI 中可见的显示板。

任务

下面是一些有关此主题的示例任务，并枚举了一些步骤。

BUI

▼ 更改显示的活动统计信息

1. 转到 "Status" (状态) > "Settings" (设置) > "Layout" (布局) 屏幕。
2. 从下拉菜单中选择要在 "Dashboard" (显示板) 中显示的统计信息。
3. 要保存您的选择，请单击 "Apply" (应用) 按钮。

▼ 更改活动阈值

1. 转到 "Status" (状态) > "Settings" (设置) > "Thresholds" (阈值) 屏幕。
2. 从下拉菜单中选择要配置的统计信息。
3. 单击 "Custom" (定制) 单选按钮。
4. 按照显示顺序定制列表中的值。有些统计信息提供 "Units" (单位) 下拉列表，所以可以选择 "Kilo"/"Mega"/"Giga"。
5. 要保存您的配置，请单击 "Apply" (应用) 按钮。

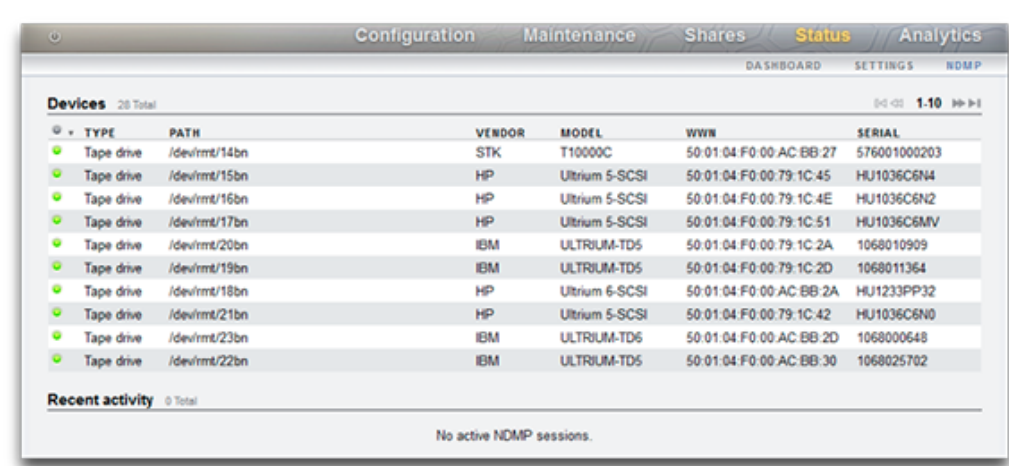
NDMP 状态

如果配置了“[NDMP 服务](#)”[\[209\]](#)且其处于活动状态，则 Status=>NDMP 页面会显示 NDMP 设备和最近的客户机活动。绿色指示灯显示设备处于联机状态，灰色指示灯显示设备处于脱机状态。

NDMP 状态 – BUI

要显示 NDMP 设备列表，请单击 "Devices"（设备）列标题。要显示有关设备的详细信息，请双击该设备。

图 2-8 NDMP 状态 BUI



NDMP 状态 – 设备

NDMP 设备在此处列出。

表 2-5 NDMP 状态 – 设备

字段	说明	示例
Type（类型）	NDMP 设备的类型	机械手、磁带机
Path（路径）	NDMP 设备的路径	/dev/rmt/14bn
Vendor（供应商）	设备供应商名称	STK
Model（型号）	设备型号名称	T1000C
WWN	World Wide Name（全球名称）	50:01:04:F0:00:AC:BB:27
Serial（序列）	设备序列号	576001000203

NDMP 状态 – 最近活动

此部分汇总近期的 NDMP 活动。

表 2-6 NDMP 状态 – 最近活动

字段	说明	示例
ID	NDMP 备份 ID	49
Active (活动)	备份当前处于活动状态	否
Remote Client (远程客户机)	NDMP 客户机地址和端口	192.168.1.219:4760
Authenticated (已验证)	显示客户机是否已完成验证	Yes (是)、No (否)
Data State (数据状态)	查看数据状态	Active (活动)、Idle (闲置) ...
Mover State (移动程序状态)	查看移动程序状态	Active (活动)、Idle (闲置) ...
Current Operation (当前操作)	当前 NDMP 操作	"Backup" (备份)、"Restore" (恢复)、"None" (无)
Progress (进度)	此备份的进度条	

NDMP Data State (NDMP 数据状态)

此字段显示备份或恢复操作的状态。可能的值如下：

- Active (活动)：数据正在备份或恢复。
- Idle (闲置)：备份或恢复尚未开始或已完成。
- Connected (已连接)：连接已建立，但是备份或恢复尚未开始。
- Halted (已停止)：备份或恢复已成功完成、已失败或已异常中止。
- Listen (侦听)：操作正在等待建立远程连接。

NDMP Mover State (NDMP 移动程序状态)

此字段显示 NDMP 设备子系统的状态。针对磁带设备的示例：

- Active (活动)：数据正在从磁带读取或正在写入磁带。
- Idle (闲置)：磁带操作尚未开始或已完成。
- Paused (已暂停)：磁带已用完或正在等待更换。
- Halted (已停止)：读取/写入操作已成功完成、已失败或已异常中止。
- Listen (侦听)：操作正在等待建立远程连接。

NDMP 状态 – CLI

NDMP 状态当前无法从 CLI 获得。

初始配置

初始配置包括以下六个部分。

- [第 4 章 网络配置](#)
- [“DNS” \[239\]](#)
- [“时间” \[243\]](#)
- 名称服务（[“NIS” \[223\]](#)、[“LDAP” \[225\]](#)、[“Active Directory” \[228\]](#)）。
- [第 5 章 存储配置](#)
- [“注册和支持” \[246\]](#)

先决条件

将在首次打开系统电源并建立连接后执行初始系统配置，如[“安装”](#)中所述。

注 - 请注意，执行群集初始配置的选项仅在 BUI 中提供。如果选择了此选项，请先阅读[第 10 章 群集配置](#)了解成功的群集设置所需的其他详细步骤，再开始初始配置。请特别注意[“网络的群集注意事项” \[160\]](#)一节。也可以通过以下过程先将支持群集的设备配置为单机运行，稍后再重新配置为在群集中运行。

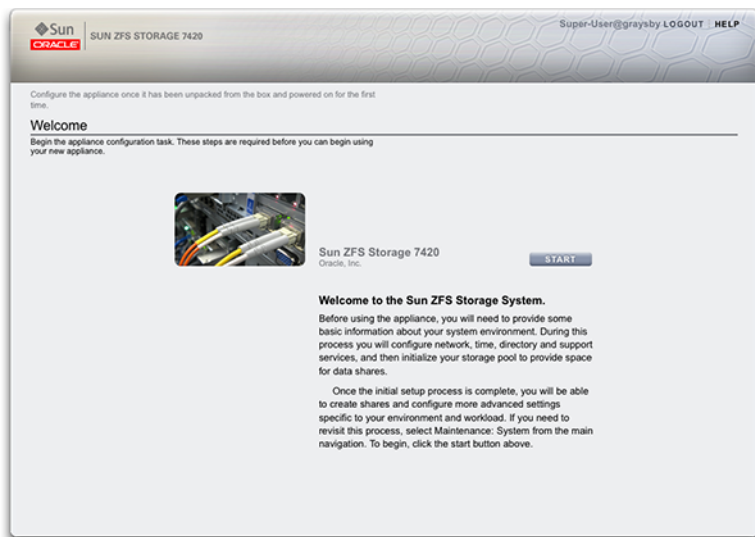
使用 BUI 执行初始配置

初始配置将配置网络连接、多个客户机网络服务以及用于单机运行的存储池布局。完成后，设备即可供使用，但不会配置供远程客户机访问的任何共享资源。要创建共享资源或重新访问设置，请参见[第 12 章 共享资源、项目和模式](#)。

以后，通过单击《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“系统”屏幕上的“INITIAL SETUP”（初始设置）按钮或在 CLI 中进入 maintenance system setup 上下文，可重复此初始配置。

BUI 初始配置是首选方法，并且为每个初始配置步骤提供了一个屏幕。

图 3-1 ZFSSA 欢迎页



▼ 执行初始配置

1. 要启动初始配置，请在欢迎页面上单击 "Start"（开始）。
2. 要在各个页面上提交您的更改并转至下一个屏幕，请单击 "Commit"（提交）。
3. 要转至上一个屏幕，请使用箭头按钮。

配置管理端口

所有单机控制器都必须至少拥有一个配置为管理接口的 NIC 端口。在 BUI 中选择 "Allow Admin"（允许管理）选项可在端口 215 上启用 BUI 连接，在 ssh 端口 22 上启用 CLI 连接。

所安装的全部群集在每个控制器上都必须至少拥有一个配置为管理接口的 NIC 端口，如上所述。此外，每个控制器上的 NIC 实例编号必须是唯一的。

使用 CLI 执行初始配置

可以使用 CLI 执行初始配置部分。每个步骤开始时都会输出其帮助（通过键入 `help` 可以再次输出帮助）。使用 `done` 命令可完成每个步骤。

使用在《[Oracle ZFS Storage Appliance 安装指南](#)》中的“安装”期间提供的密码登录：

```
caji console login: root
Password:
Last login: Sun Oct 19 02:55:31 on console
```

To setup your system, you will be taken through a series of steps; as the setup process advances to each step, the help message for that step will be displayed.

Press any key to begin initial configuration ...

在此示例中，检查了从 DHCP 服务器获取的现有设置，然后键入 `done` 接受了这些设置。要在此时定制数据，请进入每个上下文（数据链路、设备和接口），并键入 `help` 来查看该上下文可用的操作。有关其他文档，请参见[第 4 章 网络配置](#)一节。如果您要配置群集，请特别注意[“网络的群集注意事项” \[160\]](#)一节。

```
aksh: starting configuration with "net" ...
```

Configure Networking. Configure the appliance network interfaces. The first network interface has been configured for you, using the settings you provided at the serial console.

Subcommands that are valid in this context:

```
datalinks      => Manage datalinks

devices        => Manage devices

interfaces     => Manage interfaces

help [topic]   => Get context-sensitive help. If [topic] is specified,
                  it must be one of "builtins", "commands", "general",
                  "help" or "script".

show           => Show information pertinent to the current context

abort          => Abort this task (potentially resulting in a
                  misconfigured system)

done           => Finish operating on "net"
```

```
caji:maintenance system setup net> devices show
Devices:
```

DEVICE	UP	MAC	SPEED
igb0	true	0:14:4f:8d:59:aa	1000 Mbit/s
igb1	false	0:14:4f:8d:59:ab	0 Mbit/s

```

        igb2 false    0:14:4f:8d:59:ac    0 Mbit/s
        igb3 false    0:14:4f:8d:59:ad    0 Mbit/s

caji:maintenance system setup net> datalinks show
Datalinks:

    DATALINK CLASS      LINKS      LABEL
    igb0 device         igb0       Untitled Datalink

caji:maintenance system setup net> interfaces show
Interfaces:

    INTERFACE STATE CLASS LINKS      ADDR5      LABEL
    igb0 up     ip    igb0      192.168.2.80/22    Untitled Interface

caji:maintenance system setup net> done

```

有关其他与 DNS 相关的文档，请参阅[“DNS” \[239\]](#)一节。

Configure DNS. Configure the Domain Name Service.

Subcommands that are valid in this context:

```

    help [topic]      => Get context-sensitive help. If [topic] is specified,
                        it must be one of "builtins", "commands", "general",
                        "help", "script" or "properties".

    show              => Show information pertinent to the current context

    commit            => Commit current state, including any changes

    abort             => Abort this task (potentially resulting in a
                        misconfigured system)

    done              => Finish operating on "dns"

    get [prop]        => Get value for property [prop]. ("help properties"
                        for valid properties.) If [prop] is not specified,
                        returns values for all properties.

    set [prop]        => Set property [prop] to [value]. ("help properties"
                        for valid properties.) For properties taking list
                        values, [value] should be a comma-separated list of
                        values.

```

```

caji:maintenance system setup dns> show
Properties:
    <status> = online
    domain = sun.com
    servers = 192.168.1.4

caji:maintenance system setup dns> set domain=sf.fishworks.com
    domain = sf.fishworks.com (uncommitted)
caji:maintenance system setup dns> set servers=192.168.1.5
    servers = 192.168.1.5 (uncommitted)
caji:maintenance system setup dns> commit
caji:maintenance system setup dns> done

```



```
aksh: done with "dns", advancing configuration to "ntp" ...
```

配置网络时间协议 (Network Time Protocol, NTP) 以同步设备时钟。有关其他文档，请参见[“NTP” \[243\]](#)一节。

Configure Time. Configure the Network Time Protocol.

Subcommands that are valid in this context:

```

help [topic]      => Get context-sensitive help. If [topic] is specified,
                    it must be one of "builtins", "commands", "general",
                    "help", "script" or "properties".

show              => Show information pertinent to the current context

commit           => Commit current state, including any changes

abort            => Abort this task (potentially resulting in a
                    misconfigured system)

done             => Finish operating on "ntp"

enable           => Enable the ntp service

disable          => Disable the ntp service

get [prop]        => Get value for property [prop]. ("help properties"
                    for valid properties.) If [prop] is not specified,
                    returns values for all properties.

set [prop]        => Set property [prop] to [value]. ("help properties"
                    for valid properties.) For properties taking list
                    values, [value] should be a comma-separated list of
                    values.
```

```

caji:maintenance system setup ntp> set servers=0.pool.ntp.org
                    servers = 0.pool.ntp.org (uncommitted)
caji:maintenance system setup ntp> commit
caji:maintenance system setup ntp> done
aksh: done with "ntp", advancing configuration to "directory" ...
```

有关其他文档，请参阅[“NIS” \[223\]](#)、[“LDAP” \[225\]](#)和[“Active Directory” \[228\]](#)各节。

Configure Name Services. Configure directory services for users and groups. You can configure and enable each directory service independently, and you can configure more than one directory service.

Subcommands that are valid in this context:

```

nis              => Configure NIS

ldap             => Configure LDAP

ad              => Configure Active Directory

help [topic]     => Get context-sensitive help. If [topic] is specified,
```

```

                                it must be one of "builtins", "commands", "general",
                                "help" or "script".

show                            => Show information pertinent to the current context

abort                          => Abort this task (potentially resulting in a
                                misconfigured system)

done                            => Finish operating on "directory"

caji:maintenance system setup directory> nis
caji:maintenance system setup directory nis> show
Properties:
                                <status> = online
                                domain = sun.com
                                broadcast = true
                                ypservers =

caji:maintenance system setup directory nis> set domain=fishworks
                                domain = fishworks (uncommitted)
caji:maintenance system setup directory nis> commit
caji:maintenance system setup directory nis> done
caji:maintenance system setup directory> done
aksh: done with "directory", advancing configuration to "support" ...

```

配置存储池，存储池的特点是提供了底层数据冗余，并提供了在所有文件系统和 LUN 之间共享的空间。有关其他文档，请参见[第 5 章 存储配置](#)一节。

Configure Storage.

Subcommands that are valid in this context:

```

help [topic]                    => Get context-sensitive help. If [topic] is specified,
                                it must be one of "builtins", "commands", "general",
                                "help", "script" or "properties".

show                            => Show information pertinent to the current context

commit                          => Commit current state, including any changes

done                            => Finish operating on "storage"

config <pool>                   => Configure the storage pool

unconfig                       => Unconfigure the storage pool

add                             => Add additional storage to the storage pool

import                          => Search for existing or destroyed pools to import

scrub <start|stop>              => Start or stop a scrub

get [prop]                      => Get value for property [prop]. ("help properties"
                                for valid properties.) If [prop] is not specified,
                                returns values for all properties.

set pool=[pool]                 => Change current pool

```

```
caji:maintenance system setup storage> show
Properties:
    pool = pool-0
    status = online
    profile = mirror
    log_profile = -
    cache_profile = -
caji:maintenance system setup storage> done
aksh: done with "storage", advancing configuration to "support" ...
```

有关远程支持配置的其他文档，请参阅（[“回拨” \[246\]](#)）。

Remote Support. Register your appliance and configure remote monitoring.

Subcommands that are valid in this context:

tags	=> Configure service tags
scrk	=> Configure phone home
help [topic]	=> Get context-sensitive help. If [topic] is specified, it must be one of "builtins", "commands", "general", "help" or "script".
show	=> Show information pertinent to the current context
abort	=> Abort this task (potentially resulting in a misconfigured system)
done	=> Finish operating on "support"

```
caji:maintenance system setup support> done
aksh: initial configuration complete!
```


网络配置

联网配置功能允许您使用物理网络端口创建各种高级联网设置，包括链路聚合、虚拟 NIC (virtual NIC, VNIC)、虚拟 LAN (virtual LAN, VLAN) 和多路径组。然后可以为这些抽象内容定义任意数量的 IPv4 和 IPv6 地址，以便用于连接系统上的各种数据服务。

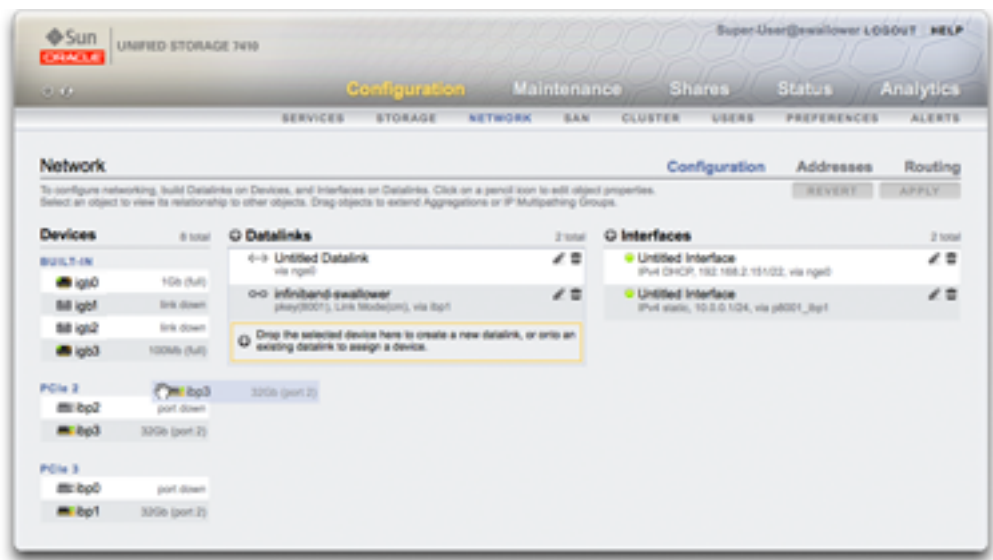
系统的网络配置有四个组成部分：

- 设备 – 物理网络端口。这些对应于您的物理网络连接或 InfiniBand 上的 IP (IP on InfiniBand, IPoIB) 分区。
- 数据链路 – 发送和接收数据包的基本结构。数据链路可以与设备（即物理网络端口）或 IB 分区一一对应，或者您可以定义由其他设备和数据链路组成的聚合、VLAN 和 VNIC 数据链路。
- 接口 – IP 配置和寻址的基本结构。每个 IP 接口都与一个数据链路关联，或者定义为包含其他接口的 IP 多路径 (IP MultiPathing, IPMP) 组。
- 路由 – IP 路由配置。这控制系统如何定向 IP 数据包。

网络配置页

在 ZFSSA 模型中，网络设备表示可用的硬件 – 它们没有可配置的设置。数据链路是第 2 层实体，必须创建它们，以便将 LACP 等设置应用到这些网络设备。接口是第 3 层实体，包含通过数据链路提供的 IP 设置。此模型将网络接口设置分为两个部分 – 数据链路对应第 2 层设置，接口对应第 3 层设置。

图 4-1 网络配置窗口



单个端口上单个 IP 地址的示例（常见配置）为：

表 4-1 示例 – 单个端口上单个 IP 地址

设备	数据链路	接口
igb0	datalink1	deimos (192.168.2.80/22)

下面的配置适用于 3 路链路聚合：

表 4-2 示例 – 适用于 3 路链路聚合的配置

设备	数据链路	接口
igb1、igb2、igb3	aggr1 (LACP 聚合)	phobos (192.168.2.81/22)

数据链路实体（我们命名为 "aggr1"）以可配置的方式（LACP 聚合策略）将网络设备组合在一起。接口实体（我们命名为 "phobos"）提供可配置的 IP 地址设置，这些设置通过数据链路在网上提供。网络设备（系统命名为 "igb1"、"igb2"、...）没有直接的设置。需要数据链路才能完成网络配置，无论它们是否将特定设置应用到网络设备。

设备

它们由系统创建，表示可用的网络或 InfiniBand 端口。它们没有自己的配置设置。

数据链路

它们管理设备，并由接口使用。它们支持：

- LACP – 链路聚合控制协议，用于捆绑多台网络设备以用作一台设备。这可以改进性能（带宽翻倍）和可靠性（能够实现网络端口容错），但是必须将设备连接到支持并为那些端口启用了 LACP 的交换机。
- IB 分区 – InfiniBand 分区，用于连接到逻辑上孤立的 IB 结构域。
- VLAN – 虚拟 LAN，用于改进本地网络安全和隔离。建议使用 VLAN 来管理设备；否则，请使用 VNIC。
- VNIC – 虚拟网络接口卡，用于将单个或聚合以太网数据链路拆分成多个虚拟（以太网）数据链路。VNIC 可以选择性地以 VLAN ID 标记，还能允许在群集中共享物理网络端口。可以在下面的[“网络的群集注意事项” \[160\]](#)一节找到逐步说明。

注 - 基于 VNIC 和基于 VLAN 的数据链路不能共享同一个 VLAN ID。

IEEE802.3ad（链路聚合）标准不明确支持多个交换机之间的聚合，但是某些供应商通过专有扩展实现多交换机支持功能。如果配置有此类扩展的交换机符合 IEEE 标准且此类扩展对于最终节点透明，则设备内支持使用此交换机。如果发生错误，Oracle 支持可能需要在单交换机配置上再现此交换机。

以下数据链路设置可用：

表 4-3 数据链路设置

属性	说明
Name（名称）	使用定义的定制名称。例如：“internal”、“external”、“adminnet”等。
Speed（速度）	使用定义的速度。有效值为 auto、10、100、1000 和 10000，分别表示自动协商、强制 10M 位/秒、强制 100M 位/秒、强制 1G 位/秒和强制 10G 位/秒。速度和双工必须要么都强制为特定值，要么都设置为自动协商。并非所有的联网设备都支持强制为所有可能的速度/双工组合。强烈建议您不要禁用自动协商。但是，如果交换机禁用了自动协商，则可能需要强制速度（和双工）以确保数据链路以预期的速度在双工模式下运行。
Duplex（双工）	使用定义的传输方向。有效 CLI 值为 auto、half 和 full，分别表示自动协商、半双工和全双工。速度和

属性	说明
	双工必须要么都强制为特定值，要么都设置为自动协商。
VLAN	使用 VLAN 头。
VLAN ID	使用定义的 VLAN 标识符：对于 VNIC 可选。
VNIC	使用 VNIC。
MTU	使用定义的最大传输单元 (maximum transmission unit, MTU) 大小。默认 MTU 为 1500 字节。指定较小的 MTU (最小 1280) 可以为数据包预留空间 (例如，针对隧道协议)。指定较大的 MTU (最大 9000) 可以提高网络性能。必须为同一 LAN 上的所有系统和交换机配置选择的 MTU。设置 MTU 值并将新的网络配置提交到系统后，可以返回到网络屏幕并查看数据链路状态，以查看具体选择的 MTU 值 (单位为字节)。请注意，VLAN 或 VNIC 无法配置为大于底层数据链路的 MTU 值。
LACP Aggregation (LACP 聚合)	使用多个网络设备 LACP 聚合。
LACP Policy (LACP 策略)	使用定义的 LACP 策略来选择传出端口。L2 散列源和目标 MAC 地址；L3 使用源和目标 IP 地址；L4 使用源和目标传输级端口
LACP Mode (LACP 模式)	使用定义的 LACP 通信模式。"Active" (活动) 模式将发送和接收 LACP 消息以协商连接并监视链路状态。"Passive" (被动) 模式仅侦听 LACP 消息。"Off" (关闭) 模式将使用聚合的链路，但是不检测链路故障或交换机配置更改。有些网络交换机配置 (包括 Cisco Etherchannel) 不使用 LACP 协议：在网络中使用非 LACP 聚合时，LACP 模式应该设置为 "off" (关闭)。
LACP Timer (LACP 计时器)	对于 "Active" (活动) 模式，使用 LACP 消息之间定义的时间间隔。
IB Partition (IB 分区)	使用 IB 分区。
Partition Key (分区键)	使用底层端口设备所属的分区 (结构域)。分区键 (partition key, pkey) 可在子网管理器上查看和配置。pkey 可以在配置子网管理器之前定义，但是数据链路会保持 "down" (未连接) 状态，直到子网分区将端口 GUID 正确配置为成员。使 HCA 端口的分区成员关系与子网管理器上的 "网络 IP 多路径 (IPMP)" [74] 和 第 10 章 群集配置 规则保持一致很重要。
IB Link Mode (IB 链路模式)	使用定义的 IB 链路模式。包括两种模式："Unreliable Datagram" (不可靠数据报) 和 "Connected" (已连接)。在 "Unreliable Datagram" (不可靠数据报) 模式下，本地队列对可与任何主机上的多个其他队列对进行通信，在未经确认的情况下在 IB 层传送消息。"Unreliable Datagram" (不可靠数据报) 模式使用的 MTU 为 2044。"Connected" (已连接) 模式使用 IB 队列

属性	说明
	对，将一个本地队列专门用来与一个专用的远程队列对进行通信。"Connected"（已连接）模式使用的 MTU 为 65520，提供的吞吐量高于 "Unreliable Datagram"（不可靠数据报）模式。

网络接口

网络接口通过数据链路配置 IP 地址。支持以下各项：

- IPv4 和 IPv6 协议。
- IPMP – IP 多路径，用于通过允许 IP 地址自动从发生故障的数据链路迁移到正常运行的数据链路来改进网络可靠性。

以下接口设置可用：

表 4-4 接口设置

属性	说明
Name（名称）	接口的定制名称
Allow Administration（允许管理）	允许通过此接口连接到设备管理 BUI 或 CLI。如果网络环境包括一个单独的管理网络，则可以仅针对管理网络启用此设置以提高安全性。
Enable Interface（启用接口）	启用此接口以用于 IP 通信。如果禁用某个接口，则设备不再通过它发送或接收 IP 流量，或者利用其上配置的任何 IP 地址。目前，禁用 IPMP 组中的活动 IP 接口不会激活备用接口。
IPv4 Configure with（IPv4 配置方式）	要么选择 "Static Address List"（静态地址列表）手动输入，要么选择 "DHCP" 动态请求
IPv4 Address/Mask（IPv4 地址/掩码）	以 CIDR 表示法表示的一个或多个 IPv4 地址 (192.168.1.1/24)
IPv6 Configure with（IPv6 配置方式）	要么选择 "Static Address List"（静态地址列表）手动输入，要么选择 "IPv6 AutoConfiguration"（IPv6 自动配置）以使用自动生成的链路本地地址（如果 IPv6 路由器响应，则为站点本地地址）
IPv6 Address/Mask（IPv6 地址/掩码）	以 CIDR 表示法表示的一个或多个 IPv6 地址 (1080::800:200C:417A/32)
IP MultiPathing Group（IP 多路径组）	配置 IP 多路径，在其中可使用数据链路池实现冗余

网络 IP 多路径 (IPMP)

IP 多路径组用于提供发生 IP 接口故障（例如物理线路断开连接或网络设备与其交换机之间的连接发生故障）时和系统及其网络网关之间发生路径故障时仍然可用的 IP 地址。系统通过监视 IP 接口的底层数据链路以获得链路接通和链路断开通知，或者可以选择使用可以为组中每个 IP 接口分配的测试地址进行探测来检测故障，如下所述。IPMP 组中可以具有任意数量的 IP 接口，只要它们都位于相同的链路（LAN、IB 分区或 VLAN）即可，而且可以为 IPMP 组分配任意数量的高可用性地址。

IPMP 组中的每个 IP 接口要么指定为 `<i>"active"`（活动）`</i>`，要么指定为 `<i>"standby"`（备用）`</i>`：

- **Active（活动）**：IP 接口将用于发送和接收数据，只要 IPMP 确定其可以正常工作即可。
- **Standby（备用）**：IP 接口仅在活动接口（或之前激活的备用接口）停止工作时用于发送和接收数据。

可以配置多个活动和备用 IP 接口，但是每个 IPMP 组都必须至少配置有一个活动 IP 接口。IPMP 力求根据需要激活任意数量的备用接口以保持配置的活动接口数量。例如，如果某个 IPMP 组配置有两个活动接口和两个备用接口，并且所有接口都可以正常工作，则仅使用两个活动接口发送和接收数据。如果某个活动接口发生故障，则将激活其中一个备用接口。如果另一个活动接口也发生故障（或已激活的备用接口发生故障），则将激活第二个备用接口。如果活动接口后来得到修复，则将重新停用备用接口。

IP 接口故障可由基于链路的检测或基于探测的检测（即配置了测试地址）发现。

如果在 IP 接口上启用了基于探测的故障检测，则系统将确定要动态探测哪些目标系统。首先，将扫描路由表以查找与 IP 接口的测试地址位于相同子网上的网关（路由器），最多可以选择五个。如果在相同的子网上未找到网关，则系统将发送一个多播 ICMP 探测（对于 IPv4，发送到 224.0.0.1；对于 IPv6，发送到 ff02::1）并选择相同子网上响应的前五个系统。因此，对于使用 IPMP 的网络故障检测和修复，应该确保每个链路上至少有一个相邻节点或默认网关响应 ICMP 回显请求。IPMP 同时适用于 IPv4 和 IPv6 地址配置。对于 IPv6，接口的链路本地地址将用作测试地址。

注 - 当配置用于响应 ICMP 回显请求的 IPMP 测试地址所在的子网上没有系统（而非群集对等设备）时，请勿使用基于探测器的故障检测。

系统将以循环模式探测所选的目标系统。如果五次连续的探测都未回应，则认为 IP 接口已发生故障。相反，如果十次连续的探测都得到了回应，则系统将认为之前发生故障的 IP 接口已修复。可以从“IPMP” [242] 屏幕设置系统的 IPMP 探测故障检测时间。该时间间接控制探测速率和修复间隔 - 例如，10 秒的故障检测时间意味着系统将以大约两秒的间隔发送探测，并且系统将需要 20 秒的时间检测基于探测的接口修复。虽然无法直接控制系统所选的目标系统，但是可以通过路由表进行间接控制。

系统将监视路由表并根据需要自动调整其选择的目标系统。例如，如果系统使用多播发现的目标，但是后来添加了路由，该路由有一个网关与 IP 接口的测试地址位于相同的子

网上，则系统将自动切换到探测网关。类似地，如果探测多播发现的目标，则系统将定期刷新其选定的目标组（例如，因为某些之前选定的目标变得无响应）。

此处[“网络 IP 多路径 \(IPMP\)” \[74\]](#)中有建立 IPMP 组的分步式说明。

有关专用本地接口的信息，请参阅[第 10 章 群集配置](#)。

网络性能和可用性

IPMP 和链路聚合是设备中提供的不同技术，用于提高网络性能和保持网络可用性。一般情况下，部署链路聚合可获取更好的网络性能，使用 IPMP 可确保高可用性。这两种技术互补，可以一起部署，取长补短，综合提高网络性能和可用性。

在链路聚合中，传入通信流量在构成聚合的多个链路上分配。因此，由于安装了更多 NIC 以将链路添加到聚合，提高了网络性能。IPMP 的通信使用 IPMP 接口的数据地址，因为它们绑定到了可用的活动接口。举例来说，如果所有数据通信只在两个 IP 地址之间流动，但不一定在同一连接上，则添加多个 NIC 不会借助 IPMP 提高性能，因为只有两个 IP 地址保持可用。

性能会受到在给定设备的数据链路上配置的 VNIC/VLAN 数量以及使用 VLAN ID 的影响。在某个给定设备上配置多个 VNIC 可能会影响该设备的所有数据链路的性能，使性能最多降低百分之五，即使未在使用 VNIC 也是如此。如果在某个给定数据链路上配置了超过八个 VNIC/VLAN，性能可能会显著下降。此外，如果数据链路使用 VLAN ID，该设备的所有数据链路性能可能会再下降百分之五。

网络路由配置

系统提供了一个 IP 路由表，它是路由表条目的集合。当需要将某个 IP 数据包发送到给定目标时，系统选择其目标与数据包目标地址最为匹配的路由条目（根据系统的多宿主策略 – 请参见下文）。然后它使用该路由条目中的信息确定通过哪个 IP 接口发送数据包，以及（如果目标无法直接到达）要使用的下一中继站网关。如果没有路由条目与该目标匹配，则将丢弃该数据包。如果有多个路由条目最匹配（并且无法通过多宿主策略确定优先顺序），则系统会按连接将负荷分摊给这些条目。

系统不会充当路由器的角色。

网络路由条目

路由表由路由条目组成，每个路由条目均包含以下字段：

表 4-5 路由条目字段

字段	说明	示例
Destination (目标)	可与路由匹配的 IP 目标地址范围 (采用 CIDR 表示法)	192.168.0.0/22
Gateway (网关)	要将数据包发送至的下一中继站 (IP 地址) ("系统"路由除外 - 请参见下文)	192.168.2.80
Family (族)	Internet 协议	IPv4、IPv6
Type (类型)	路由的来源	dhcp、static (静态)、system (系统)
Interface (接口)	借以发送数据包的 IP 接口	igb0

"destination" (目标) 字段为 `0.0.0.0/0` 的路由条目与任何数据包都匹配 (如果没有其他更精确匹配的路由), 因此称为 "default" (默认) 路由。在 BUI 中, 默认路由与非默认路由通过另一属性进行区分:

表 4-6 区分默认路由与非默认路由

Kind (种类)	路由类型	Default (默认值)、Network (网络)

如上所述, 给定数据包将通过在路由条目的 "interface" (接口) 字段中指定的 IP 接口发送。如果指定了 IPMP 接口, 则将按连接随机选择 IPMP 组中的活动 IP 接口之一, 并在选择的 IP 接口后来变得不可用时自动刷新。相反, 如果给定 IP 接口是 IPMP 组的一部分, 则它无法在 "interface" (接口) 字段中指定, 因为此类路由不具有高可用性。

路由条目有多种不同的来源, 来源通过 "type" (类型) 字段进行标识。尽管路由条目的来源与系统如何使用该路由条目无关, 但是来源却控制是否以及如何编辑或删除该路由条目。系统支持以下类型的路由:

表 4-7 支持的路由类型

类型	说明
Static (静态)	由设备管理员创建和管理
system (系统)	由设备在启用 IP 接口的过程中自动创建。系统路由针对设备可以直接到达的每个 IP 子网创建。由于这些路由都可以直接到达, 所以 "gateway" (网关) 字段标识设备在该子网上的 IP 地址。
DHCP	在启用配置为使用 DHCP 的 IP 接口的过程中由设备自动创建。DHCP 路由针对 DHCP 服务器提供的每个默认路由创建。

类型	说明
Dynamic（动态）	由设备通过 RIP 和 RIPng 动态路由协议（如果已启用）自动创建。

还有一种类型标识当前不能使用的静态路由：

表 4-8 不可用的静态路由类型

Inactive（非活动）	之前创建并且与已禁用或脱机 IP 接口关联的静态路由。

网络路由属性

表 4-9 路由属性

属性	说明
Multihoming model（多宿主模式）	在同时启用多个 IP 接口时控制接受和传输 IP 数据包的系统策略。允许的值包括 "loose"（宽松）（默认）、"adaptive"（自适应）和 "strict"（严格）。请参见下文的论述。

如果系统配置有多个 IP 接口，则可能有多个等效的路由通往给定目标，强制系统选择在哪个 IP 接口上发送数据包。类似地，数据包可能会先到达一个 IP 接口，但目标却是另一 IP 接口上托管的 IP 地址。在此类情况下，系统行为由所选的多宿主策略确定。支持三种策略：

表 4-10 多宿主策略

策略	说明
Loose（宽松）	不强制 IP 数据包与用于发送或接收该数据包的 IP 接口之间建立任何绑定：1) 只要 IP 数据包的目标 IP 地址位于设备上，即可通过某个 IP 接口接收该数据包。2) IP 数据包将通过与该 IP 数据包的目标地址最明确匹配的路由绑定的 IP 接口传输，而不考虑该 IP 接口上托管的 IP 地址。如果没有符合条件的路由存在，则丢弃该数据包。
Adaptive（自适应）	与 loose（宽松）相同，但是优先选择具有与数据包源 IP 地址位于相同子网上的网关地址的路由：1) 只要 IP 数据包的目标 IP 地址位于设备上，即可通过某个 IP 接口接收该数据包。2) IP 数据包将通过与该 IP 数据包的目标地址最明确匹配的路由绑定的 IP 接口传输。如果多个路由同样明确，则优先选择具有与数据包源地址位于相同子网上的网关地址的路由。如果没有符合条件的路由存在，则丢弃该数据包。

策略	说明
Strict (严格)	要求 IP 数据包与用于发送或接收该数据包的 IP 接口之间建立严格绑定：1) 只要 IP 数据包的目标 IP 地址位于某个 IP 接口上，即可通过该 IP 接口接收该数据包。2) 只有 IP 数据包的源 IP 地址位于某个 IP 接口上，才能通过该 IP 接口传输该数据包。为了强制实施此要求，当针对可用的路由进行匹配时，设备将忽略网关地址与数据包的源地址位于不同子网上的任何路由。如果没有符合条件的路由存在，则丢弃该数据包。

当选择多宿主策略时，一个重要注意事项就是是否有任何设备的 IP 接口专用于管理（例如，专供 BUI 访问），因而通过单独的管理网络进行访问。具体而言，如果创建一个默认路由以提供对管理网络的远程访问，并创建一个单独的默认路由以提供对存储协议的远程访问，则默认系统策略 "loose"（宽松）可能会导致管理默认路由用于存储通信。通过将策略切换到 "adaptive"（自适应）或 "strict"（严格），设备将在选择路由进行回复的过程中考虑与请求关联的 IP 地址。如果在相同的 IP 接口上找不到路由，则 "adaptive"（自适应）策略将导致系统使用任意可用的路由，而 "strict"（严格）策略将导致系统丢弃该数据包。

使用 BUI 进行网络配置

当使用 BUI 重新配置联网时，系统会尽力保持浏览器的当前网络连接。但是，某些网络配置更改（例如删除浏览器连接的特定地址）将不可避免地导致浏览器断开其连接。出于此原因，建议分配特定的 IP 地址和网络设备供管理员使用，并且始终配置该地址。如果需要，还可以通过串行控制台从 CLI 执行特别复杂的网络重新配置任务。

以下图标在 "Configuration"（配置）> "Network"（网络）部分中使用：

表 4-11 网络配置图标

图标	说明
	添加新的数据链路/接口/路由
	编辑数据链路/接口/路由设置
	编辑功能已禁用
	销毁数据链路/接口/路由
	销毁功能已禁用
	拖放图标

图标	说明
	已连接的网络端口
	具有 I/O 活动的已连接网络端口
	已断开连接的网络端口（链路断开，线缆问题？）
	活动 InfiniBand 端口
	具有 I/O 活动的活动 InfiniBand 端口
	非活动 InfiniBand 端口（down、init 或 arm 状态）
	InfiniBand 分区设备正常工作
	InfiniBand 分区设备发生故障（子网管理器问题）
	网络数据链路
	网络数据链路 VLAN 或 VNIC
	网络数据链路聚合
	网络数据链路聚合 VLAN 或 VNIC
	网络数据链路 IB 分区
	正在使用接口发送和接收数据包（正常工作或者已降级）
	接口已由用户禁用
	接口处于脱机状态（归群集对等设备所有）
	接口发生故障或者配置了重复的 IP 地址

右上方是 "Configuration"（配置）、"Addresses"（地址）和 "Routing"（路由）的本地导航，显示备用配置视图。

网络配置页

默认显示 "Configuration"（配置）页，并列出 "Devices"（设备）、"Datalinks"（数据链路）和 "Interfaces"（接口），以及一些用于管理的按钮。将鼠标悬停在某个条目上还会显示一个  图标，单击任何条目均可突出显示与其关联的其他组件。

"Devices"（设备）列表在右侧显示链路状态，以及反映网络端口状态的图标。如果端口显示已断开连接，请检查是否已正确将它们插入网络。

要在网络设备上配置 IP 地址，请首先创建一个数据链路，然后创建一个要使用该数据链路的接口。 图标完成这两项任务，创建过程中将显示 "Datalink"（数据链路）和 "Interface"（接口）属性对话框。

可通过多种方式配置网络接口。尝试单击某个设备对应的  图标，然后将其拖到数据链路表中。再将数据链路拖到接口表中。还可执行其他移动操作。这有助于完成复杂的配置，配置过程中会突出显示有效的移动。

网络地址

此页显示当前网络配置的摘要表，具有如下字段：

表 4-12 当前网络配置摘要

字段	说明	示例
Network Datalink（网络数据链路）	数据链路名称和详细信息摘要	datalink1（通过 igb0）
Network Interface（网络接口）	接口名称和详细信息摘要	IPv4 DHCP，通过 datalink1
Network Addresses（网络地址）	此接口托管的地址	192.168.2.80/22
Host Names（主机名）	网络地址对应的已解析主机名	caji.sf.example.com

网络路由页

此页提供 IP 路由表和关联属性的配置，如上文所述。默认情况下，路由表中的所有条目都会显示，但是可以使用子导航栏按类型筛选该表。

要检查特定路由，请在 CLI 中使用 traceroute。

```
zfssa-source:> traceroute 10.80.198.102
traceroute: Warning: Multiple interfaces found; using 10.80.198.101 @ igb3
traceroute to 10.80.198.102 (10.80.198.102), 30 hops max, 40 byte packets
 1 10.80.198.1 (10.80.198.1) 6.490 ms 0.924 ms 0.834 ms
 2 10.80.198.102 (10.80.198.102) 0.152 ms 0.118 ms 0.099 ms
zfssa-target:> traceroute 10.80.198.101
traceroute: Warning: Multiple interfaces found; using 10.80.198.102 @ igb3
traceroute to 10.80.198.101 (10.80.198.101), 30 hops max, 40 byte packets
 1 10.80.198.1 (10.80.198.1) 1.031 ms 0.905 ms 0.769 ms
 2 10.80.198.101 (10.80.198.101) 0.158 ms 0.111 ms 0.109 ms
```


使用 CLI 进行网络配置

网络配置位于 configuration net 下，它具有用于 devices、datalinks、interfaces 和 routing 的子命令。可将 show 命令与各项子命令结合使用来显示当前配置：

```
caji:> configuration net
caji:configuration net> devices show
Devices:
```

DEVICE	UP	SPEED	MAC
igb0	true	1000 Mbit/s	0:14:4f:9a:b9:0
igb1	true	1000 Mbit/s	0:14:4f:9a:b9:1
igb2	true	1000 Mbit/s	0:14:4f:9a:b8:fe
igb3	true	1000 Mbit/s	0:14:4f:9a:b8:ff

```
caji:configuration net> datalinks show
Datalinks:
```

DATALINK	CLASS	LINKS	LABEL
igb0	device	igb0	datalink1

```
caji:configuration net> interfaces show
Interfaces:
```

INTERFACE	STATE	CLASS	LINKS	ADDRS	LABEL
igb0	up	ip	igb0	192.168.2.80/22	caji

```
caji:configuration net> routing show
Properties:
```

```
    multihoming = loose
```

```
Routes:
```

ROUTE	DESTINATION	GATEWAY	INTERFACE	TYPE
route-000	0.0.0.0/0	192.168.1.1	igb0	dhcp
route-001	192.168.0.0/22	192.168.2.142	igb0	system

在每个部分键入 help 可查看用于创建和配置数据链路、接口和路由的相关命令。此上下文中有效的子命令：

```
help [topic]      => Get context-sensitive help. If [topic] is specified,
                    it must be one of "builtins", "commands", "general",
                    "help", "script" or "properties".

show              => Show information pertinent to the current context

commit           => Commit current state, including any changes

abort            => Abort creation of "vnic"

done             => Finish operating on "vnic"

get [prop]       => Get value for property [prop]. ("help properties"
                    for valid properties.) If [prop] is not specified,
                    returns values for all properties.
```

```

set [prop]          => Set property [prop] to [value]. ("help properties"
                    for valid properties.) For properties taking list
                    values, [value] should be a comma-separated list of
                    values.

available           => Get values that can be assigned to the links
                    parameter when creating a network component.

```

available 命令可用于在创建网络组件时查看可分配给 links 参数哪些值。下面显示了 CLI 命令 available 的输出：

```

caji:configuration net datalinks> device
caji:configuration net datalinks device (uncommitted)> available
igb7,igb6

caji:configuration net datalinks> vnic
caji:configuration net datalinks vnic (uncommitted)> available
igb5,igb4,aggr2,aggr1

caji:configuration net datalinks> vlan
caji:configuration net datalinks vlan (uncommitted)> available
igb5,igb4,aggr2,aggr1

caji:configuration net datalinks> aggregation
caji:configuration net datalinks aggregation (uncommitted)> available
igb7,igb6

caji:configuration net interfaces> ip
caji:configuration net interfaces ip (uncommitted)> available
aggr2,aggr1

caji:configuration net interfaces> ipmp
caji:configuration net interfaces ipmp (uncommitted)> available
vnic4,vnic3,igb5,igb4

```

以下示例说明了如何使用 device 命令创建数据链路，以及如何使用 ip 命令创建接口：

```

caji:configuration net> datalinks
caji:configuration net datalinks> device
caji:configuration net datalinks device (uncommitted)> set links=igb1
links = igb1 (uncommitted)
caji:configuration net datalinks device (uncommitted)> set label=datalink2
label = datalink2 (uncommitted)
caji:configuration net datalinks device (uncommitted)> set mtu=9000
mtu = 9000 (uncommitted)
caji:configuration net datalinks device (uncommitted)> commit
caji:configuration net datalinks> show
Datalinks:

    DATALINK CLASS      LINKS      LABEL
    igb0 device        igb0       datalink1
    igb1 device        igb1       datalink2

caji:configuration net datalinks> cd ..
caji:configuration net> interfaces
caji:configuration net interfaces> ip

```

```

caji:configuration net interfaces ip (uncommitted)> set label="caji2"
      label = caji2 (uncommitted)
caji:configuration net interfaces ip (uncommitted)> set links=igb1
      links = igb1 (uncommitted)
caji:configuration net interfaces ip (uncommitted)> set v4addr=10.0.1.1/8
      v4addr = 10.0.1.1/8 (uncommitted)
caji:configuration net interfaces ip (uncommitted)> commit
caji:configuration net interfaces> show
Interfaces:

```

INTERFACE	STATE	CLASS	LINKS	ADDRS	LABEL
igb0	up	ip	igb0	192.168.2.80/22	caji
igb1	up	ip	igb1	10.0.1.1/8	caji2

以下示例说明了如何通过新 igb1 IP 接口上的 10.0.1.2 创建默认路由：

```

caji:configuration net routing> create
caji:configuration net route (uncommitted)> set family=IPv4
      family = IPv4 (uncommitted)
caji:configuration net route (uncommitted)> set destination=0.0.0.0
      destination = 0.0.0.0 (uncommitted)
caji:configuration net route (uncommitted)> set mask=0
      mask = 0 (uncommitted)
caji:configuration net route (uncommitted)> set interface=igb1
      interface = igb1 (uncommitted)
caji:configuration net route (uncommitted)> set gateway=10.0.1.2
      gateway = 10.0.1.2 (uncommitted)
caji:configuration net route (uncommitted)> commit

```

使用 BUI 执行网络配置任务

▼ 创建单个端口接口

1. 单击 "Datalinks" (数据链路)  图标。
2. (可选) 设置名称并选择定制 MTU 单选按钮 (在文本框中键入 9000)。
3. 从 "Devices" (设备) 列表选择一个设备。
4. 单击 "APPLY" (应用)。该数据链路将显示在 "Datalinks" (数据链路) 列表中。
5. 单击 "Interface" (接口)  图标。
6. 设置所需的属性，并选择之前创建的数据链路。
7. 单击 "APPLY" (应用)。该接口将显示在 "Interfaces" (接口) 列表中。

8. 正在运行设备的网络配置尚未发生更改。完成接口配置后，单击顶部的 "APPLY"（应用）以提交配置。

▼ 修改接口

1. 单击数据链路或接口上的编辑图标。
2. 将设置更改为所需的值。
3. 单击对话框上的 "APPLY"（应用）。
4. 单击页面顶部的 "APPLY"（应用）以提交配置。

▼ 通过拖放创建单个端口接口

1. 将鼠标悬停在设备上并单击拖放图标 (⊕)。
2. 将其拖到 "Datalinks"（数据链路）列表，然后释放。
3. （可选）设置名称和大型 MTU。
4. 单击 "APPLY"（应用）。
5. 接下来，将数据链路拖到 "Interfaces" 列表中。
6. 设置所需的属性，然后单击 "APPLY"（应用）。
7. 单击屏幕顶部的 "APPLY"（应用）以提交配置。

▼ 创建 LACP 聚合链路接口

1. 单击 "Datalinks"（数据链路） 图标。
2. （可选）设置数据链路名称。
3. 选择 "LACP Aggregation"（LACP 聚合）。
4. 从 "Devices"（设备）列表中选择两个或更多设备，然后单击 "APPLY"（应用）。

5. 单击 "Interfaces" (接口)  图标。
6. 设置所需的属性，从 "Datalinks" (数据链路) 列表中选择聚合链路，然后单击 "APPLY" (应用)。
7. 单击顶部的 "APPLY" (应用) 以提交配置。

▼ 创建使用基于探测和链路状态故障检测的 IPMP 组

当配置用于响应 ICMP 回显请求的 IPMP 测试地址所在的子网上没有系统（而非群集对等设备）时，请勿使用基于探测器的故障检测。

1. 创建一个或多个“底层”IP 接口以用作 IPMP 组的组件。每个接口都必须有一个要用作探测源的 IP 地址（参见上面创建单个端口接口的单独任务）。
2. 单击 "Interface" (接口)  图标。
3. （可选）更改接口的名称。
4. 单击 "IP MultiPathing Group" (IP 多路径组) 复选框。
5. 单击 "Use IPv4 Protocol" (使用 IPv4 协议) 或/和 "Use IPv6 Protocol" (使用 IPv6 协议)，然后指定 IPMP 接口的 IP 地址。
6. 从 "Interfaces" (接口) 列表中选择第一步中创建的接口。
7. 根据需要每个选择的接口设置为 "Active" (活动) 或 "Standby" (备用)。
8. 单击 "APPLY" (应用)。

▼ 创建仅使用链路状态故障检测的 IPMP 组

1. 创建一个或多个具有 IP 地址 0.0.0.0/8 的“底层”IP 接口，以用作 IPMP 组的组件（参见上面创建单个端口接口的单独任务）。
2. 单击 "Interface" (接口)  图标。
3. （可选）更改接口的名称。
4. 单击 "IP MultiPathing Group" (IP 多路径组) 复选框。

5. 单击 "Use IPv4 Protocol" (使用 IPv4 协议) 或/和 "Use IPv6 Protocol" (使用 IPv6 协议) , 然后指定 IPMP 接口的 IP 地址。
6. 从 "Interfaces" (接口) 列表中选择第一步中创建的接口。
7. 根据需要将每个选择的接口设置为 "Active" (活动) 或 "Standby" (备用) 。
8. 单击 "APPLY" (应用) 。

▼ 扩展 LACP 聚合

1. 将鼠标悬停在 "Devices" (设备) 列表中的设备上。
2. 单击  图标, 将设备拖到聚合数据链路上, 然后释放。
3. 单击页面顶部的 "APPLY" (应用) 以提交该配置。

▼ 扩展 IPMP 组

1. 将鼠标悬停在 "Interfaces" (接口) 列表中的接口上。
2. 单击  图标, 将设备拖到 IPMP 接口上, 然后释放。
3. 单击页面顶部的 "APPLY" (应用) 以提交该配置。

▼ 创建 InfiniBand 分区数据链路和接口

1. 单击 "Datalinks" (数据链路)  图标。
2. (可选) 设置名称。
3. 单击 "IB Partition" (IB 分区) 复选框
4. 从 "Partition Devices" (分区设备) 列表选择一个设备。
5. 单击 "APPLY" (应用) 。新的分区数据链路将显示在 "Datalinks" (数据链路) 列表中。

6. 单击 "Interface" (接口)  图标。
7. 设置所需的属性，并选择之前创建的数据链路。
8. 单击 "APPLY" (应用)。该接口将显示在 "Interfaces" (接口) 列表中。
9. 正在运行设备的网络配置尚未发生更改。完成接口配置后，单击顶部的 "APPLY" (应用) 以提交配置。

▼ 为群集控制器创建不带 VLAN ID 的 VNIC

本例针对一半网络端口处于备用状态的主动-主动配置。本任务在设备数据链路上创建一个 IP 接口，并将其分配给一个机头。在同一数据链路的基础上构建一个 VNIC，在 VNIC 的基础上配置一个 IP 接口，并将其分配给另一个机头。在一个给定数据链路上配置一个而非多个 VNIC 可以确保高峰性能。流量流经与一个机头上底层活动端口关联的电缆以及与另一个机头上底层备用端口关联的电缆。因此，闲置备用端口可以与 VNIC 一起使用。

1. 当群集处于 AKCS_CLUSTERED 状态时，请单击数据链路  图标。
2. (可选) 设置名称和 MTU。
3. 从 "Devices" (设备) 列表选择一个设备，然后单击 "APPLY" (应用)。该数据链路将显示在 "Datalinks" (数据链路) 列表中。
4. 单击 "Interface" (接口)  图标。
5. 设置所需的属性，并选择之前创建的数据链路，然后单击 "APPLY" (应用)。该接口将显示在 "Interfaces" (接口) 列表中。
6. 单击 "Datalinks" (数据链路)  图标。
7. 选中 "VNIC" 复选框，选择性地设置名称和 MTU (等于或小于第 2 步中的值)，然后单击 "APPLY" (应用)。新 VNIC 数据链路会显示在 "Datalinks" (数据链路) 列表中。
8. 单击 "Interface" (接口)  图标。
9. 设置所需的属性，选择之前创建的 VNIC 数据链路，然后单击 APPLY (应用)。该接口将显示在 "Interfaces" (接口) 列表中。

10. 正在运行设备的网络配置尚未发生更改。完成接口配置后，单击顶部的 "APPLY"（应用）以提交配置。
11. 单击 "Cluster"（群集）选项卡。新建的两个接口与默认所有者显示在 "Resource"（资源）部分。
12. 使用 "Owner"（所有者）下拉列表将两个接口中的一个分配给另一个机头，然后单击 "APPLY"（应用）。

▼ 为群集控制器创建带有相同 VLAN ID 的 VNIC

本例针对一半网络端口处于备用状态的主动-主动配置。本任务在同一个设备数据链路的基础上创建两个带有相同 VLAN ID 的 VNIC。每个 VNIC 配置有一个接口，将每个接口分配给不同的机头。流量流经与一个机头上底层活动端口关联的电缆以及与另一个机头上底层备用端口关联的电缆。因此，闲置备用端口可以与 VNIC 一起使用。

1. 当群集处于 AKCS_CLUSTERED 状态时，请单击数据链路  图标。
2. 选中 "VNIC" 复选框，选择性地设置名称和 MTU，设置 VLAN ID，从 "Devices"（设备）列表中选择一设备，然后单击 "APPLY"（应用）。新 VNIC 数据链路会显示在 "Datalinks"（数据链路）列表中。
3. 单击 "Interface"（接口） 图标。
4. 设置所需的属性，选择之前创建的 VNIC 数据链路，然后单击 APPLY（应用）。该接口将显示在 "Interfaces"（接口）列表中。
5. 按照第 1 步和第 2 步中的说明再创建一个带有相同设备 ID 和 VLAN ID 的 VNIC，然后按照第 3 步和第 4 步中的说明为其创建一个接口。
6. 正在运行设备的网络配置尚未发生更改。完成接口配置后，单击顶部的 "APPLY"（应用）以提交配置。
7. 单击 "Cluster"（群集）选项卡。新建的两个接口与默认所有者显示在 "Resource"（资源）部分。
8. 使用 "Owner"（所有者）下拉列表将两个接口中的一个分配给另一个机头，然后单击 "APPLY"（应用）。

▼ 添加静态路由

1. 转至 "Configuration" (配置) > "Network" (网络) > "Routing" (路由)
2. 单击添加图标。
3. 根据之前的说明填写属性。
4. 单击 "ADD" (添加)。新的路由将显示在表中。

▼ 删除静态路由

1. 转至 "Configuration" (配置) > "Network" (网络) > "Routing" (路由)
2. 将鼠标悬停在路由条目上，然后单击右侧的垃圾箱图标。

使用 CLI 执行网络配置任务

▼ 添加静态路由

1. 转至 `configuration net routing`。
2. 输入 `create`。
3. 键入 `show` 以列出所需属性，然后 `set` 每个属性。
4. 输入 `commit`。

▼ 删除静态路由

1. 转至 `configuration net routing`。
2. 键入 `show` 以列出路由以及路由名称 (例如 `route-002`)。
3. 输入 `destroy` 后跟路由名称。

▼ 将多宿主属性更改为 strict

1. 转至 `configuration net routing`
2. 输入 `set multihoming=strict`
3. 输入 `commit`

存储配置

存储配置在具有各自底层数据冗余特性的池中，存储提供的空间在所有文件系统和 LUN 之间共享。有关存储池如何与单个文件系统或 LUN 关联的更多信息，请参见[“共享资源” \[262\]](#)一节。

每个节点都可以具有任意数量的池，并且可以为群集中的每个池单独分配所有权。虽然支持任意数量的池，但是不建议创建具有相同冗余特性且归同一群集机头所有的多个池。这样做会导致性能不佳、资源分配欠佳、产生虚拟存储分区，并增加管理复杂性。仅当需要完全不同的冗余或性能特性时才建议在同一主机上配置多个池，例如镜像池和 RAID-Z 池。由于能够按共享资源控制对日志和高速缓存设备的访问，因此建议的操作模式是单个池。

池可以通过配置新池或导入现有池进行创建。导入现有池仅用于导入之前在 Sun Storage 7000 设备上配置的池，并且在意外进行了重新配置、在机头节点之间移动池或发生灾难性机头故障时非常有用。

当将原始存储分配到池时，请记住填满池将导致性能明显降低，特别是在写入共享资源或 LUN 时。这些影响通常在超过池容量的 80% 时开始显现，并在超过池容量的 90% 时变得比较明显。因此，可以通过多置备大约 20% 的容量来获得最佳结果。[“共享资源” \[262\]](#) UI 可用于确定当前使用的空间量。

存储配置文件

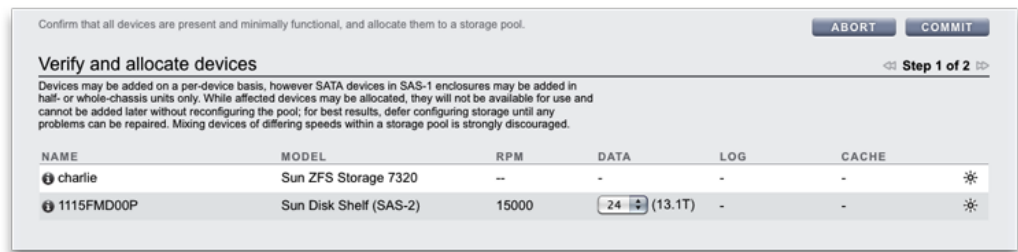
图 5-1 存储配置文件



此操作配置存储池。在 BUI 中，可以通过单击池列表旁边的  按钮完成此任务，单击时将提示您输入新池的名称。在 CLI 中，可以通过 config 命令完成此任务，该命令将池的名称用作参数。

在任务开始后，存储配置分为两个不同的阶段：验证和配置。

图 5-2 验证和分配设备



存储配置规则和准则

为了获得最佳性能，请牢记以下事项：

规则 1 – 包含在机头节点或 JBOD 中的所有“数据”磁盘都必须具有相同的旋转速度（介质旋转速率）。ZFSSA 软件将检测错误配置，并针对该情况生成错误。

建议 1 – 由于不可预测的性能问题，请避免在同一池中混用不同的磁盘旋转速度。

建议 2 – 为了获得最佳的性能，请勿在同一 SAS 结构（HBA 连接）上混用具有不同磁盘旋转速度的 JBOD。这样混用后虽然可以正常运行，但是可能会降低原本运行速度较快的设备的性能。

建议 3 – 当配置包含不同容量数据磁盘的存储池时，ZFS 在某些情况下会将最小容量磁盘的大小用于存储池中的部分或全部磁盘，从而使预期的整体容量变小。使用的大小将取决于设备的存储配置文件、布局和组合。避免在同一池中混用不同的磁盘容量。

存储验证

验证可以确保所有存储都已连接且正常运行。所有存储设备必须都已连接并正常运行，然后才能对其进行分配。如果为池分配缺少或发生故障的设备，稍后将无法添加缺少或发生故障的设备。

在不带连接的存储的系统中，默认情况下会分配所有可用驱动器。在可扩展的系统中，磁盘机框与机头节点一起显示在列表中，并且可以在每个磁盘机框中控制分配。根据机头节点或磁盘架的型号，这项操作的方式可能不同。

您可以选择以下内容：

- Device size（设备大小） – 按逻辑大小筛选数据设备。默认情况下，“Any”显示所有可用的数据设备。

- Data devices（数据设备） – 显示所有可用的数据设备或选定设备大小的可用数量。

默认情况下分配的磁盘数量取决于以下几项：

- The maximum number available（可用的最大数量） – 当连接的存储仅包含大小相同和旋转速度相同的设备时，或者从多个磁盘大小中选择一个大小时
- None（无） – 当连接的存储中混用旋转速度不同的设备时。

注意：为了提供稳定的性能特征，强烈建议池仅包含大小相同且旋转速度相同的设备。

SAS-2 系统上的存储分配

虽然所有机箱中的驱动器都可以单独分配，但是当从 JBOD 分配磁盘时应小心谨慎以保证最佳池配置。总的来说，池越少，每个池的磁盘越多越好，因为这样可以简化管理并提高整体可用容量的百分比。

虽然系统可以根据需要以任意增量分配存储，但是建议每项分配至少包含所有 JBOD 中的 8 个磁盘，而且越多越理想。

数据配置文件配置

验证完成后，下一步是选择反映设置的 RAS 和性能目标的存储配置文件。系统将根据可用存储提供一组可能的配置文件。下表列出了所有可能的配置文件及其说明。

表 5-1 数据配置文件配置

数据配置文件	说明
双重奇偶校验选项	
Triple mirrored（三重镜像）	数据进行三重镜像，以实现高度可靠且高性能的系统（例如，用于重要数据库的存储）。此配置适合需要最高性能和可用性的情况。与双重镜像相比，三重镜像在每个存储块中增加了一个 IOPS 并针对故障实现了更高级别的保护。注意：不带扩展存储的控制器不得配置有三重镜像。
Double parity RAID（双重奇偶校验 RAID）	采用 RAID，其中每个存储条包含两个奇偶校验磁盘。与三重镜像一样，可实现高可用性，因为数据在任意两个磁盘发生故障的情况下仍然可用。"Double parity RAID"（双重奇偶校验 RAID）与镜像选项相比是容量更高的选项，并且适用于高吞吐量的序列访问工作负荷（例如备份）或使用低随机读取组件存储大量数据的情况。

数据配置文件	说明
单奇偶校验选项	
Mirrored（镜像）	数据进行镜像，这样会将容量减半，但是可实现高度可靠且高性能的系统。建议在空间充足且对性能具有较高要求时（例如，数据库存储）使用。
Single parity RAID, narrow stripes（单奇偶校验 RAID，窄存储条）	采用 RAID，其中每个存储条包含三个数据磁盘和一个奇偶校验磁盘。对于可接受单奇偶校验保护的情况，单奇偶校验 RAID 与简单镜像相比是容量更高的选项。此更高容量需要通过比镜像选项更低的随即读取功能进行平衡。对于具有中等随机读取组件的非关键型应用程序，可以考虑单奇偶校验 RAID。对于纯粹的流式工作负荷，应优先使用 "Double parity RAID"（双重奇偶校验 RAID）选项，因为它具有更高的容量和更大的吞吐量。
其他	
Striped（带区）	数据在磁盘之间进行条带化处理，并且没有冗余。虽然这能同时最大限度地提高性能和容量，但是单个磁盘故障即会导致数据丢失。建议不要使用此配置。对于纯粹的流式工作负荷，请考虑使用 "Double parity RAID"（双重奇偶校验 RAID）。
Triple parity RAID, wide stripes（三重奇偶校验 RAID，宽存储条）	采用 RAID，其中每个存储条包含三个奇偶校验磁盘。这是除了条带化数据以外容量最高的选项。由于存储条较宽且随机 I/O 性能较低，因此在一个或多个驱动器出现故障后重新同步数据的时间会明显延长。与其他的 RAID 配置一样，高速缓存的存在可以缓解对读取性能的影响。一般不建议使用此配置。

对于可扩展系统，某些配置文件可能提供 "NSPF" 选项。这代表 "no single point of failure"（无单点故障），表示数据在镜像或 RAID 存储条中排列，因此严重的 JBOD 故障不会造成数据丢失。请注意，系统几乎已在所有组件中配置了冗余。每个 JBOD 都有冗余路径、冗余控制器以及冗余电源和风扇。NSPF 只能保护数据不受磁盘底板故障（一个几乎完全被动的组件）或重大不当管理行为（将指向一个 JBOD 的两个路径断开）的影响。一般来说，采用 NSPF 将导致容量较低，因为它对存储条宽度具有较为严格的要求。

日志设备只能使用条带化或镜像配置文件进行配置。由于日志设备仅在节点出现故障时使用，因此只有在两个设备同时出现故障且节点随后立即重新引导时，未镜像日志的数据才会丢失。这种事件发生的可能性极小，因为它要求在非常短的时间段内出现两个并发设备故障和节点故障，但是镜像日志设备可以有效地避免这种情况。

注意：当不同尺寸的日志设备位于不同的机箱时， 仅可创建条带化日志配置文件。

热备件分配为池总大小的百分比，并独立于所选的配置文件（条带化除外，因为它不支持热备件）。由于热备件针对每个存储配置步骤进行分配，所以将存储配置为一个整体将比以较小的增量添加存储高效得多。

在群集中，高速缓存设备仅可用于导入了存储池的节点。在群集中，可以同时两个节点上的高速缓存设备配置为相同池的一部分。要执行此操作，请接管被动节点上的池，然后添加存储并选择高速缓存设备。这样便可以随时一次性配置一半的全局高速缓存设备。虽然高速缓存设备上的数据会在故障转移时丢失，但是新的高速缓存设备可以用于新的节点。

注意：早期的软件版本支持采用宽存储条的双重奇偶校验。这已经由采用宽存储条的三重奇偶校验所取代，因为它能显著提高可靠性。在早期软件版本中配置为采用宽存储条的双重奇偶校验的池继续受到支持，但是新配置的或重新配置的池无法选择该选项。

导入现有存储池

通过使用此功能，您可以导入现有存储池，以及任何意外取消配置的池。此功能可在恢复出厂设置或执行维修操作后用于恢复用户数据。导入池需要迭代所有连接的存储设备并搜索所有现有的状态。这可能需要相当长的时间，并且在此期间无法执行任何其他存储配置活动。要在 BUI 中导入池，请单击存储配置屏幕中的 "IMPORT"（导入）按钮。要在 CLI 中导入池，请使用 "import" 命令。

搜索阶段完成后，系统将向您提供可用池的列表，包括一些标识特征。如果存储已销毁或不完整，则该池不可导入。与存储配置不同，池名称不在开始的时候指定，而是在选择池的时候指定。默认情况下，使用之前的池名称，但是可以通过在 BUI 中单击该名称或在 CLI 中设置 "name" 属性来更改池名称。

添加其他存储

使用此操作可将其他存储添加到现有池。验证步骤与初始配置期间的验证步骤相同。存储必须使用最初用于配置池的相同配置文件进行添加。如果存储不足，无法使用当前配置文件配置系统，则可以舍弃一些属性。例如，将单个 JBOD 添加到双重奇偶校验 RAID-Z NSPF 配置将无法保留 NSPF 特性。但是，仍然可以添加 JBOD 并在 JBOD 中创建 RAID 存储条，在此过程中将舍弃 NSPF。

取消配置存储

这会删除任何活动的文件系统和 LUN 并取消配置存储池，让原始数据存储可在将来进行存储配置时使用。此过程可以通过导入未配置的存储池进行撤消，前提是尚未将原始存储用作活动存储池的一部分。

存储池清理

这会启动存储池清理过程，该过程将检验所有的内容，看其是否有错误。如果找到无论是通过清理还是常规操作都无法恢复的任何错误，BUI 将显示受影响的文件。如果需要，也可以停止清理。

使用 BUI 配置存储

▼ 配置存储池

可通过两种途径完成此任务：设备的初始配置过程，或 "Configuration"（配置）-> "Storage"（存储）屏幕。

1. 单击存储池列表上方的  按钮
2. 输入存储池的名称
3. 在 "Allocate and verify storage"（分配和验证存储）屏幕中，为存储池配置 JBOD 分配。JBOD 分配可以为 "none"（无）、"half"（半）或 "all"（全部）。如果未检测到 JBOD，请检查 JBOD 布线和电源。
4. 单击 "COMMIT"（提交）。
5. 在 "Configure Added Storage"（配置添加的存储）屏幕上，选择所需的数据配置文件。每个都按照可用性、性能和容量进行评级，帮助找到适合您业务需求的最佳配置。
6. 单击 "COMMIT"（提交）。

▼ 将高速缓存设备添加到现有池中

1. 将新的 Readzilla 或 Logzilla 设备安装到第一个可用插槽中。有关插槽位置，请参见 [《Oracle ZFS Storage Appliance 安装指南》](#) 中的“概述”。
2. 在 BUI 中，转至 "Configuration"（配置）> "Storage"（存储）。
3. 在 "Available Pools"（可用池）列表中，选择设备要添加到的池。确保此池处于联机状态。

4. 单击 "Add" (添加) 按钮以将设备添加到此池。
5. 选择要添加到此池的设备，并单击 "Commit" (提交)。
6. 选择日志配置文件 (如果适用)，并单击 "Commit" (提交)。

使用 CLI 配置存储

▼ 将高速缓存设备添加到现有池中

1. 将新的 Readzilla 或 Logzilla 设备安装到第一个可用插槽中。有关插槽位置，请参见 [《Oracle ZFS Storage Appliance 安装指南》](#) 中的“概述”。
2. 在命令行中输入：
3. `poc:> configuration storage`
4. 指定此设备要添加到的池：
5. `poc:configuration storage (pool_2)> set pool=pool_2`
6. `pool = pool_2`
7. `poc:configuration storage (pool_2)> add`
8. 一则消息提醒您验证是否正确安装此设备。请注意，强烈建议不要混用设备类型和速度。
9. 显示此池的设备信息：
10. `poc:configuration storage (pool_2) verify> show`
11. **ID STATUS ALLOCATION DATA LOG CACHE RPM**
12. **0 ok custom 0 0 0/4 1.86T**
13. **1 ok custom 0 0/2 34G 0 15000**
14. **2 ok custom 0 0/2 34G 0 15000**

15. 指定要使用的磁盘机框和 Logzilla 或 Readzilla 数量。在下例中，1-log=1 从第一个磁盘机框分配一个 Logzilla。
16. `poc:configuration storage (pool_2) verify> set 1-log=1`
17. `1-log = 1`
18. 注意：值 "1-log=2" 将从第一个磁盘机框分配两个 Logzilla。
19. 此示例将从第一个磁盘机框分配两个 Readzilla。
20. `poc:configuration storage (pool_2) verify> set 1-cache=1`
21. `1-cache = 1`
22. 输入 done。
23. `poc:configuration storage (pool_2) verify> done`
24. 注意：如果您在一个池中添加了奇数数量的 Logzilla 设备或者一个池尚无配置文件，可输入 `set log_profile=log_mirror` 设置日志配置文件。
25. 输入 show 以显示配置文件。
26. `poc:configuration storage (pool_2) config> show`
27. `:`
28. `PROFILE CAPCTY NSPF DESCRIPTION`
29. `log_profile = log_stripe 17G no Striped log`
30. 输入 done 以完成此任务：
31. `poc:configuration storage (pool_2) config> done`
32. `poc:configuration storage (pool_2)>`

存储区域网络配置

SAN 配置页面允许您将设备连接到 SAN（Storage Area Network，存储区域网络）。SAN 由三个基本组件组成：

- 访问网络存储的客户机
- 提供网络存储的存储设备
- 将客户机链接到存储的网络

不管网络上使用的是何种协议，这三个组件都保持不变。在某些情况下，网络甚至可能是启动器和目标之间的一条线缆，但是大多数情况下，涉及某些类型的交换。

SAN 目标和启动器

目标和启动器按协议进行配置。有关详细信息，请参阅特定协议（“[SAN 光纤通道](#)” [105]、“[iSCSI](#)” [189]或“[SRP](#)” [122]）的相关文档。

SAN 目标和启动器组

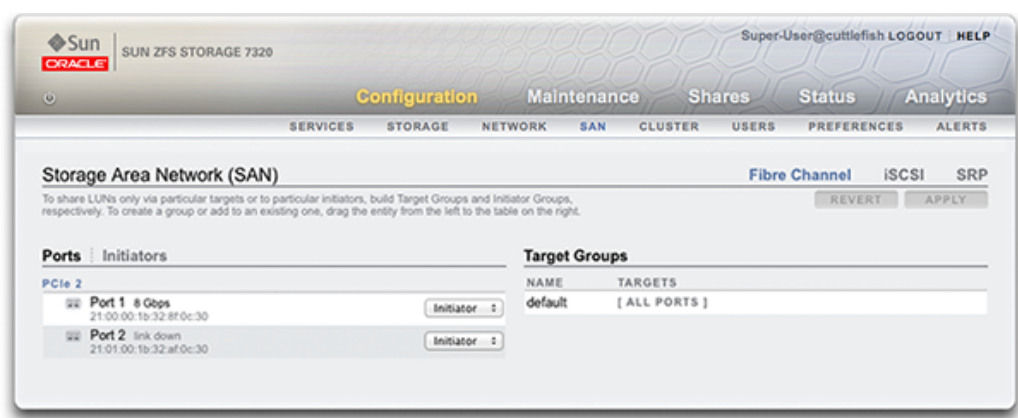
目标和启动器组定义可以与 LUN 关联的目标和启动器集。与某个目标组关联的 LUN 仅能通过该组中的目标进行查看。如果某个 LUN 未显式与目标组关联，则它位于默认目标组中，并能够通过所有的目标进行访问，不管协议是什么。同样，只有 LUN 所属的一个或多个组中的启动器可以看到它。如果某个 LUN 未显式与启动器组关联，则它位于默认启动器组中，并可以通过所有启动器访问。虽然使用默认启动器组对于评估可能非常有用，但是不建议这么做，因为这可能导致将 LUN 暴露给不必要或冲突的启动器。

为了避免一个启动器属于多个组时 LUN 可能发生冲突，要在将组与 LUN 关联之前在所有组中配置启动器。

使用 BUI 配置 SAN

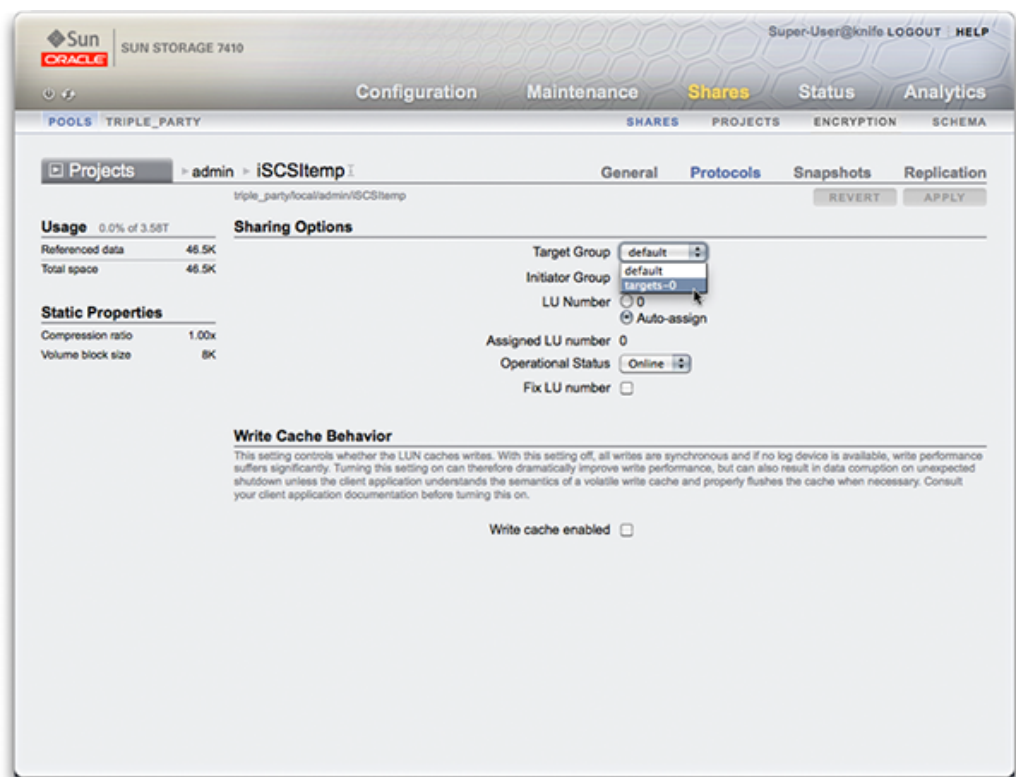
要配置目标，请转至 "Configuration"（配置）> "SAN BUI" 页面，使用 "Fibre Channel"（光纤通道）、"iSCSI" 和 "SRP" 进行导航，然后配置 "Ports"（端口）、"Initiator"（启动器）和 "Target Groups"（目标组）控件。

图 6-1 SAN BUI 页面



要关联 LUN，请转至 "Shares"（共享资源）> "Shares"（共享资源）> "Protocols"（协议）页面，然后配置 "Target Group"（目标组）和 "Initiator Group"（启动器组）控件。

图 6-2 关联 LUN



使用 CLI 配置 SAN

可使用 CLI 的 configuration san 上下文按协议类型对目标和启动器执行操作。然后，可使用 shares CLI 上下文创建 LUN，并将其与目标和启动器组关联。

SAN 术语

要配置设备以在 SAN 上运行，您应理解一些基本的 SAN 术语：

表 6-1 SAN 术语

术语	说明
SCSI 目标	SCSI 目标是一个存储系统端点，它提供一个服务以处理来自启动器的 SCSI 命令和 I/O 请求。SCSI 目标由存储系统管理员创建，并通过唯一寻址方法进行标识。在配置后，SCSI 目标将包含零个或多个逻辑单元。
SCSI 启动器	SCSI 启动器是能够启动 SCSI 会话、发送 SCSI 命令和 I/O 请求的应用程序或生产系统端点。SCSI 启动器也通过唯一寻址方法进行标识（请参见 SCSI 目标）。
逻辑单元	“逻辑单元”这个术语用于描述存储系统中的一种组件。分配唯一编号后，就产生了所谓的逻辑单元号 (Logical Unit Number, LUN)。一个高度可配置的存储系统可以包含许多 LUN。当这些 LUN 与一个或多个 SCSI 目标关联时，便形成了唯一的 SCSI 设备（可以通过一个或多个 SCSI 启动器访问）。
iSCSI	Internet SCSI，一种协议，用于通过 IP 网络共享基于 SCSI 的存储。
iSER	RDMA 的 iSCSI 扩展 (iSCSI Extension for RDMA)，一种协议，用于通过提供 RDMA 服务的网络（即 InfiniBand）映射 iSCSI 协议。iSER 协议由 iSCSI 子系统通过透明的方式进行选择，前提是已正确配置 IB 硬件。在 CLI 和 BUI 中，所有支持 iSER 的组件（目标和启动器）都作为 iSCSI 组件进行管理。
FC	光纤通道 (Fibre Channel)，一种协议，用于通过由光缆、FC 交换机和 HBA 组成的存储区域网络 (Storage Area Network, SAN) 共享基于 SCSI 的存储。
SRP	SCSI RDMA 协议 (SCSI RDMA Protocol)，该协议通过提供 RDMA 服务的网络（即 InfiniBand）共享基于 SCSI 的存储。
IQN	iSCSI 限定名称 (iSCSI Qualified Name)，iSCSI 网络中设备的唯一标识符。iSCSI 对 IQN 使用 iqn.date.authority:uniqueid 格式。例如，设备可以使用以下 IQN 来标识其中一个 iSCSI 目标：iqn.1986-03.com.sun:02:c7824a5b-f3ea-6038-c79d-ca443337d92c。此名称表示这是一个由 1986 年 3 月注册的公司所构建的 iSCSI 设备。命名权威机构只不过是颠倒的公司 DNS 名称，在此示例中为 "com.sun"。后面的整个字符串是 Oracle 用于标识该目标的唯一 ID。
目标门户	使用 iSCSI 协议时，目标门户指的是 IP 地址和 TCP 端口号的唯一组合，通过该组合启动器可以与目标联系。
目标门户组	使用 iSCSI 协议时，目标门户组是目标门户的集合。目标门户组以透明的方式进行管理；每个网络接

术语	说明
	口都具有与该接口的有效地址相对应的目标门户组。将目标绑定到某个接口即意味着 iSCSI 目标将使用与该接口关联的门户组。
CHAP	质询握手身份验证协议 (Challenge-Handshake Authentication Protocol)，该安全协议可以在启动器中对目标进行验证和/或在目标中对启动器进行验证。
RADIUS	代表存储节点使用中央服务器执行 CHAP 验证的系统。
目标组	一组目标。LUN 针对一个特定目标组中的所有目标导出。
启动器组	一组启动器。当启动器组与 LUN 相关联时，只有该组的启动器可以访问 LUN。
目标	一个存储系统端点，提供一个服务以处理来自启动器的 SCSI 命令和 I/O 请求。目标由存储系统管理员创建，并通过唯一寻址方法标识。在配置后，目标将包含零个或更多个逻辑单元。
启动器	能够启动 SCSI 会话、发送 SCSI 命令和 I/O 请求的应用程序或生产系统端点。启动器也通过唯一寻址方法进行标识。

每个 LUN 都有一些控制卷导出方式的属性。有关更多信息，请参见[“协议” \[290\]](#)一节。

SAN 光纤通道

光纤通道 (Fibre Channel, FC) 是几乎专门用作 SCSI 传输介质的千兆位高速网络技术。FC 是设备支持的多个块协议之一；要通过 FC 共享 LUN，设备必须配备一个或多个可选的 FC 卡。

FC 端口目标配置

默认情况下，所有 FC 端口都配置为处于目标模式下。如果设备用于连接到磁带 SAN 进行备份，必须将一个或多个端口配置为处于启动器模式下。要将一个端口配置为处于启动器模式下，必须对设备进行重置。多个端口可以同时配置为处于启动器模式下。

每个 FC 端口都分配有一个全球名称 (World Wide Name, WWN)，并且与其他块协议一样，FC 目标可以组合成[“SAN 目标和启动器组” \[101\]](#)，以便将端口带宽专门用于特定的 LUN 或 LUN 组。将 FC 端口配置为目标后，即可检查和验证远程搜索到的端口。

有关使用 Oracle ZFS Storage Appliance 的 FC SAN 引导解决方案的详细信息，请参阅《*Implementing Fibre Channel SAN Boot with Oracle's Sun ZFS Storage Appliance*》（《使用 Oracle Sun ZFS Storage Appliance 实施光纤通道 SAN 引导》）白皮书，网址为：<http://www.oracle.com/technetwork/articles/servers-storage-admin/fbsanboot-365291.html> (<http://www.oracle.com/technetwork/articles/servers-storage-admin/fbsanboot-365291.html>)。

群集注意事项

在群集中，启动器将具有两个指向每个 LUN 的路径（或路径集）：一个路径（或路径集）指向已导入与 LUN 关联的存储的机头；另一个路径（或路径集）指向该机头的群集对等设备。第一个路径（或路径集）是活动路径；第二个路径（或路径集）是备用路径；在接管时，活动路径将变得不可用，备用路径将（在很短的时间内）变为活动路径，之后 I/O 将继续。此多路径方法称为不对称逻辑单元访问 (Asymmetric Logical Unit Access, ALUA)，当与识别 ALUA 的启动器组合使用时，可实现群集接管对于较高级别应用程序的透明化。

FC 启动器配置

启动器通过其 WWN 进行标识，并且与其他的块协议一样，可以为其创建别名。要协助为 FC 启动器创建别名，可以从搜索到的端口的 WWN 中选择一个 WWN。另外，与其他的块协议一样，启动器可以组合成组。当一个 LUN 与特定的启动器组关联时，该 LUN 仅对该组中的启动器可见。在大多数 FC SAN 中，LUN 始终与为其创建了 LUN 的系统所对应的启动器组关联。

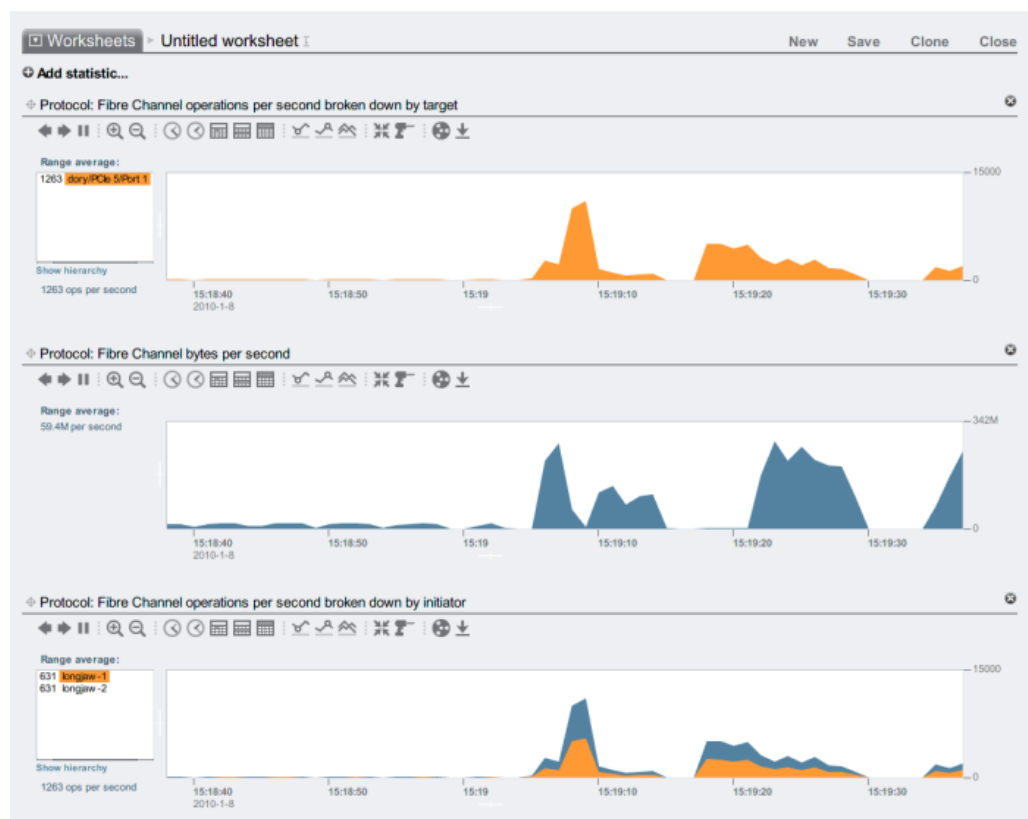
群集注意事项

设备是符合 ALUA 的阵列。要在 ALUA 环境中正确配置 FC 启动器，需要可识别 ALUA 的驱动程序，并可能需要进行特定于启动器的调整。有关详细信息，请参见 "Oracle ZFS Storage Appliance: How to set up Client Multipathing"（文档 ID 1628999.1）。

性能注意事项

可以通过《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“Analytics”监测 iSCSI 性能，在其中可按启动器、目标或 LUN 细分操作或吞吐量。

图 6-3 FC 性能



对于操作，还可以按偏移、延迟、大小和 SCSI 命令进行细分，以便不仅了解 FC 操作的内容，还了解其方式和原因。

FC 故障排除

FC 队列溢出

设备设计为使用一组全局资源为每个机头上的 LUN 提供服务。因此通常不需要限制客户机上的队列深度，因为设备的 FC 端口可以处理大量的并发请求。即使这样，还是存在极小的队列溢出可能，而队列溢出会导致 SCSI 传输错误。此类队列溢出通常与以下一个或多个原因有关：

- 前端端口过载 – 与一个 FC 端口关联的主机太多和/或通过一个 FC 端口访问的 LUN 太多
- 设备在降级模式下运行，例如在设计为主动-主动群集配置的环境中进行群集接管时

虽然队列溢出的可能性很小，但是如果愿意根据客户机限制队列深度，则可以完全消除该可能性。要确定合适的队列深度限制，应该使用目标端口的数量乘以每个端口的最大并发命令 (2048)，然后用得到的数字除以置备的 LUN 数量。要适应降级运行模式，应该对群集对等设备中的 LUN 数量求和，以确定 LUN 的数量，但是将两个群集对等设备的最小值用作目标端口的数量。例如，在主动-主动 7420 双机头群集中，如果一个机头具有 2 个 FC 端口和 100 个 LUN，而另一个机头具有 4 个 FC 端口和 28 个 LUN，则应该使用最保守的最大队列深度：两个端口乘以 2048 个命令，除以 100 个 LUN 与 28 个 LUN 的和 – 即每个 LUN 32 个命令。

最大队列深度的调整是特定于启动器的，但是在 Solaris 上，这将通过调整全局变量 `ssd_max_throttle` 来实现。

FC 链路级别问题

要对链路级别问题进行故障排除，例如光纤损坏或电缆接触不良，请查看每个 FC 端口的错误统计信息：如果任何数字明显为非零或不断增加，则可能表明出现了链路问题，应该执行链路级别诊断。

使用 BUI 配置 FC

更改 FC 端口的模式

要使用 FC 端口，请在 BUI 的 "Configuration" > "SAN" 屏幕上将其设置为 "Target"（目标）模式（使用下面屏幕抓图中显示的下拉菜单）。必须具有 root 权限才能执行此操作。请注意：在群集配置中，需要分别在每个机头节点上将端口设置为 "Target"（目标）模式。



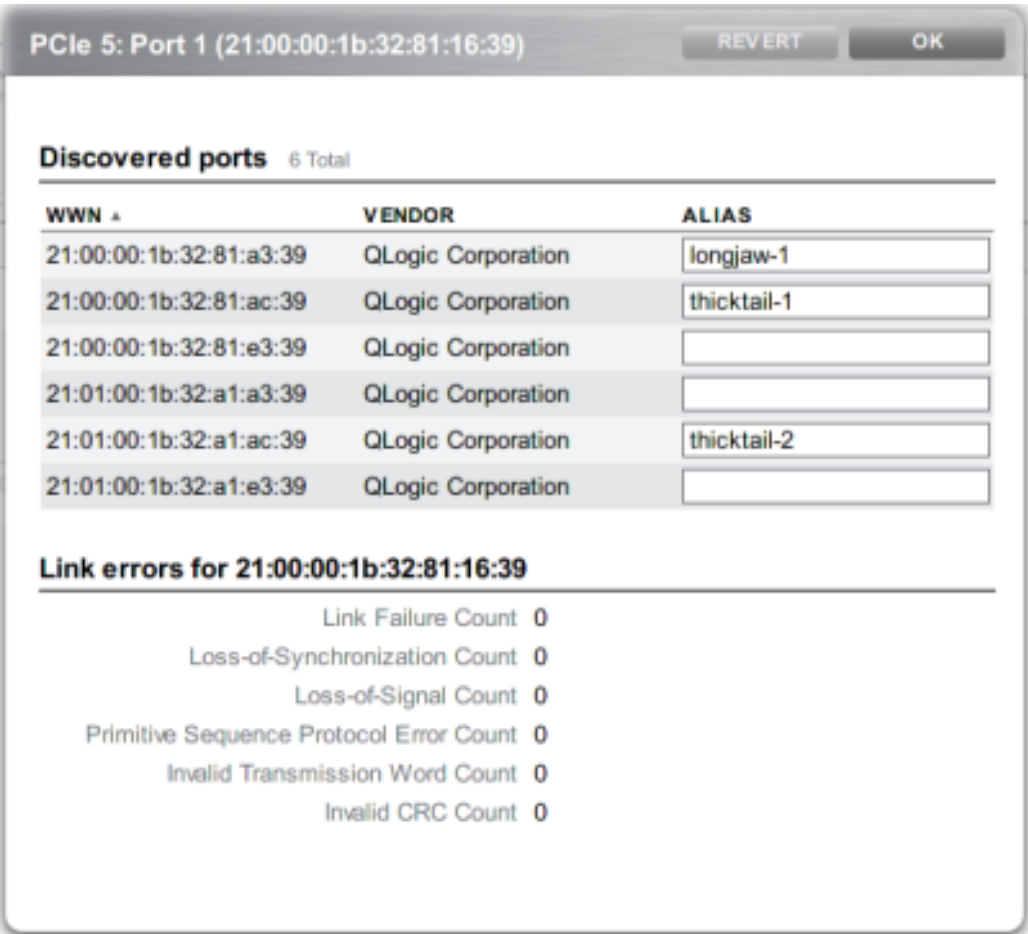
将所需的端口设置为 "Target"（目标）后，单击 "Apply"（应用）按钮。此时将显示一条确认消息，通知您设备将立即重新引导。确认您要重新引导。

设备引导完成后，活动的 FC 目标将显示  图标，并且将鼠标悬停在上面时将显示移动图标 .

查看搜索到的 FC 端口

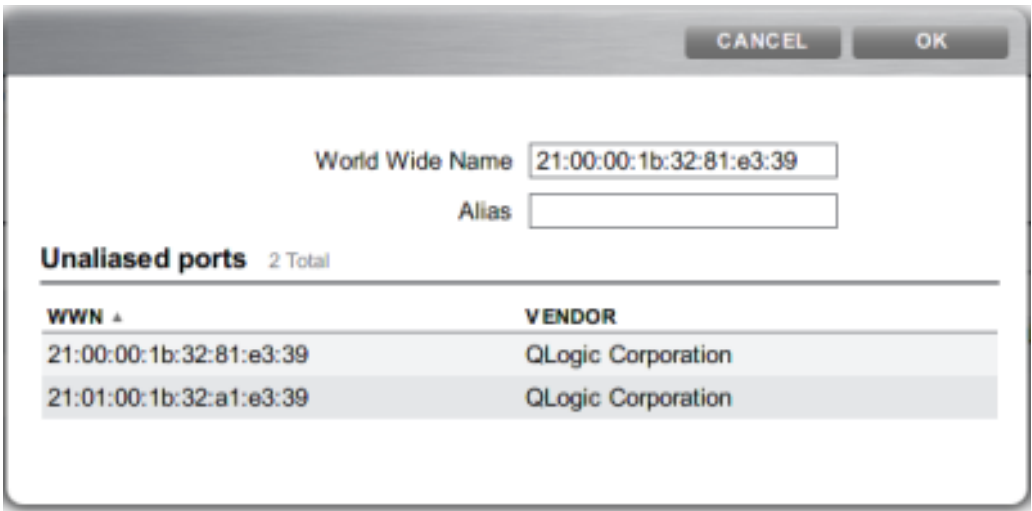
单击信息图标  可查看 "Discovered Ports"（搜索到的端口）对话框，在其中您可以对链路错误进行故障排除。在 "Discovered Ports"（搜索到的端口）对话框中，单击列表中的 WWN 可查看关联的链路错误。

图 6-4 搜索到的 FC 端口



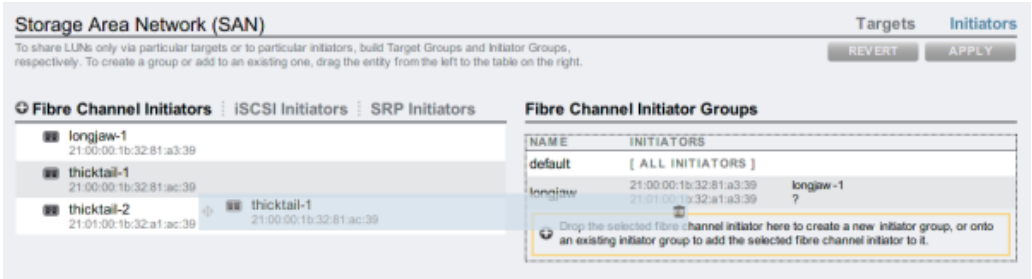
创建 FC 启动器组

可以在 "Initiators" (启动器) 屏幕中创建和管理启动器组。单击添加图标  可查看无别名的端口。单击列表中的 WWN 可在 "Alias" (别名) 字段中添加有意义的别名。



在 "Initiators" (启动器) 页面中, 通过将启动器拖到 "FC Initiator Groups" (FC 启动器组) 列表中可创建新组或将其添加到现有组中。

图 6-5 FC 启动器组列表

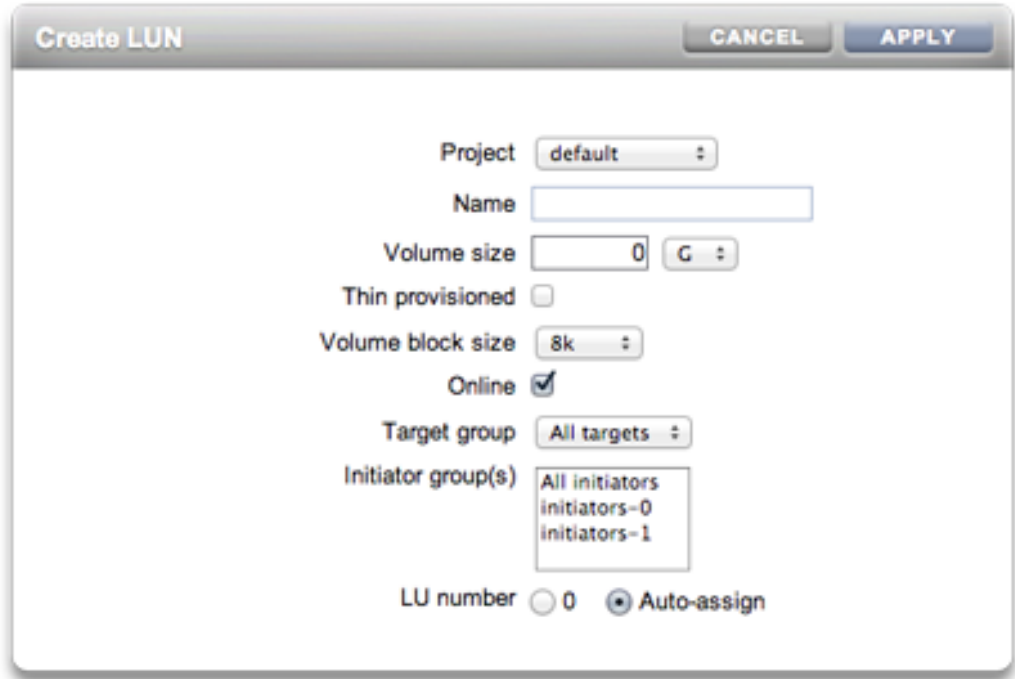


单击 "Apply" (应用) 按钮可提交新启动器组。现在, 您可以创建对客户机启动器组具有独占访问权限的 LUN。

将 LUN 与 FC 启动器组关联

要创建 LUN, 请滚动到相应启动器组上并单击添加 LUN 图标 。将显示 "Create LUN" (创建 LUN) 对话框, 其中关联的启动器组处于选中状态。设置名称和大小, 然后单击 "Apply" (应用) 将 LUN 添加到存储池。

图 6-6 将 LUN 与 FC 启动器组关联



使用 CLI 配置 FC

更改 FC 端口的模式

```
dory:configuration san fc targets> set targets="wnn.2101001B32A11639"
      targets = wnn.2101001B32A11639 (uncommitted)
dory:configuration san fc targets> commit
```

查看搜索到的 FC 端口

```
dory:configuration san fc targets> show
Properties:
      targets = wnn.2100001B32811639,wnn.2101001B32A12239
Targets:
```



```

NAME          MODE      WWN              PORT              SPEED
target-000    target    wwn.2100001B32811639  PCIe 5: Port 1    4 Gbit/s
target-001    initiator wwn.2101001B32A11639  PCIe 5: Port 2    0 Gbit/s
target-002    initiator wwn.2100001B32812239  PCIe 2: Port 1    0 Gbit/s
target-003    target    wwn.2101001B32A12239  PCIe 2: Port 2    0 Gbit/s
dory:configuration san fc targets> select target-000
dory:configuration san fc targets target-000> show
Properties:
    wwn = wwn.2100001B32811639
    port = PCIe 5: Port 1
    mode = target
    speed = 4 Gbit/s
    discovered_ports = 6
    link_failure_count = 0
    loss_of_sync_count = 0
    loss_of_signal_count = 0
    protocol_error_count = 0
    invalid_tx_word_count = 0
    invalid_crc_count = 0
Ports:
PORT      WWN              ALIAS              MANUFACTURER
port-000  wwn.2100001B3281A339  longjaw-1          QLogic Corporation
port-001  wwn.2101001B32A1A339  longjaw-2          QLogic Corporation
port-002  wwn.2100001B3281AC39  thicktail-1        QLogic Corporation
port-003  wwn.2101001B32A1AC39  thicktail-2        QLogic Corporation
port-004  wwn.2100001B3281E339  <none>              QLogic Corporation
port-005  wwn.2101001B32A1E339  <none>              QLogic Corporation

```

创建 FC 启动器组

```

dory:configuration san fc initiators> create
dory:configuration san fc initiators (uncommitted)> set name=lefteye
dory:configuration san fc initiators (uncommitted)>
    set initiators=wwn.2101001B32A1AC39,wwn.2100001B3281AC39
dory:configuration san fc initiators (uncommitted)> commit
dory:configuration san fc initiators> list
GROUP      NAME
group-001  lefteye
|
+-> INITIATORS
    wwn.2101001B32A1AC39
    wwn.2100001B3281AC39

```

将 LUN 与 FC 启动器组关联

以下示例演示如何创建名为 lefty 的 LUN，并将其与 fera 启动器组关联。

```

dory:shares default> lun lefty
dory:shares default/lefty (uncommitted)> set volsize=10
    volsize = 10 (uncommitted)
dory:shares default/lefty (uncommitted)> set initiatorgroup=fera

```

```

        initiatorgroup = default (uncommitted)
dory:shares default/lefty (uncommitted)> commit

```

编写用于创建启动器和启动器组别名的脚本

有关如何修改和使用以下示例脚本的信息，请参阅[“CLI 用法” \[35\]](#)和[“简单 CLI 脚本和批处理命令” \[35\]](#)部分。

```

script
/*
 * This script creates both aliases for initiators and initiator
 * groups, as specified by the below data structure. In this
 * particular example, there are five initiator groups, each of
 * which is associated with a single host (thicktail, longjaw, etc.),
 * and each initiator group consists of two initiators, each of which
 * is associated with one of the two ports on the FC HBA. (Note that
 * there is nothing in the code that uses this data structure that
 * assumes the number of initiators per group.)
 */
groups = {
    thicktail: {
        'thicktail-1': 'wwn.2100001b3281ac39',
        'thicktail-2': 'wwn.2101001b32a1ac39'
    },
    longjaw: {
        'longjaw-1': 'wwn.2100001b3281a339',
        'longjaw-2': 'wwn.2101001b32a1a339'
    },
    tecopa: {
        'tecopa-1': 'wwn.2100001b3281e339',
        'tecopa-2': 'wwn.2101001b32a1e339'
    },
    spinedace: {
        'spinedace-1': 'wwn.2100001b3281df39',
        'spinedace-2': 'wwn.2101001b32a1df39'
    },
    fera: {
        'fera-1': 'wwn.2100001b32817939',
        'fera-2': 'wwn.2101001b32a17939'
    }
};
for (group in groups) {
    initiators = [];
    for (initiator in groups[group]) {
        printf('Adding %s for %s ... ',
            groups[group][initiator], initiator);
        try {
            run('select alias=' + initiator);
            printf('(already exists)\n');
            run('cd ..');
        } catch (err) {
            if (err.code != EAKSH_ENTITY_BADSELECT)
                throw err;
            run('create');
        }
    }
}

```

```

        set('alias', initiator);
        set('initiator', groups[group][initiator]);
        run('commit');
        printf('done\n');
    }
    run('select alias=' + initiator);
    initiators.push(get('initiator'));
    run('cd ..');
}
printf('Creating group for %s ... ', group);
run('groups');
try {
    run('select name=' + group);
    printf('(already exists)\n');
    run('cd ..');
} catch (err) {
    if (err.code != EAKSH_ENTITY_BADSELECT)
        throw err;
    run('create');
    set('name', group);
    run('set initiators=' + initiators);
    run('commit');
    printf('done\n');
}
run('cd ..');
}

```

iSCSI

Internet SCSI 是设备所支持的多种用于共享基于 SCSI 的存储的块协议之一。

目标配置

使用 iSCSI 协议时，目标门户指的是 IP 地址和 TCP 端口号的唯一组合，通过该组合启动器可以与目标联系。

使用 iSCSI 协议时，目标门户组是目标门户的集合。目标门户组以透明的方式进行管理；每个网络接口都具有与该接口的有效地址相对应的目标门户组。将目标绑定到某个接口即意味着 iSCSI 目标将使用与该接口关联的门户组。

注意：不支持每个会话多个连接。

IQN (iSCSI Qualified Name, iSCSI 限定名称) 是 iSCSI 网络中设备的唯一标识符。iSCSI 对 IQN 使用 `iqn.date.authority:uniqueid` 格式。例如，设备可以使用以下 IQN 来标识其中一个 iSCSI 目标：`iqn.1986-03.com.sun:02:c7824a5b-f3ea-6038-c79d-ca443337d92c`。此名称表示这是一个由 1986 年 3 月注册的公司所构建的 iSCSI 设备。命名权威机构只不过是颠倒的公司 DNS 名称，在此示例中为 "com.sun"。后面的整个字符串是 Oracle 用于标识该目标的唯一 ID。

表 6-2 iSCSI 目标属性

目标属性	说明
Target IQN (目标 IQN)	此目标的 IQN。IQN 可以手动指定，也可以自动生成。
Alias (别名)	此目标的用户可读昵称。
Authentication mode (验证模式)	None、CHAP 或 RADIUS 之一。
CHAP name (CHAP 名称)	如果使用 CHAP 验证，则为 CHAP 用户名。
CHAP secret (CHAP 机密)	如果使用 CHAP 验证，则为 CHAP 密钥。
Network interfaces (网络接口)	其目标门户用于导出此目标的接口。

除了这些属性以外，BUI 还指示目标是处于联机状态还是脱机状态：

表 6-3 目标状态图标

图标	说明
	目标处于联机状态
	目标处于脱机状态

群集注意事项

在群集平台中，在相应群集节点上具有至少一个活动接口的目标将处于联机状态。将接口分配到目标时应谨慎；可能会将目标配置为使用分离的机头节点上的门户组。在这种情况下，目标在两个机头上都处于联机状态，并将根据每个机头节点所拥有的存储导出不同的 LUN。随着接管/故障恢复或所有权更改期间网络接口在群集机头之间迁移，在导入和导出各 iSCSI 目标的网络接口时这些目标将在联机和脱机状态之间进行切换。

绑定到 IPMP 接口的目标将仅通过该 IPMP 组的地址收到通告。无法通过该组的测试地址与该目标通信。绑定到基于 LACP 聚合而构建的接口的目标将使用该聚合的地址。如果将 LACP 聚合添加到 IPMP 组中，则目标将不能再使用该聚合的接口，因为该地址将成为 IPMP 测试地址。

启动器配置

iSCSI 启动器具有以下可配置的属性。

表 6-4 iSCSI 启动器属性

属性	说明
Initiator IQN (启动器 IQN)	此启动器的 IQN。
Alias (别名)	此启动器的用户可读昵称。
Use CHAP (使用 CHAP)	启用或禁用 CHAP 验证
CHAP name (CHAP 名称)	如果使用 CHAP 验证, 则为 CHAP 用户名。
CHAP secret (CHAP 机密)	如果使用 CHAP 验证, 则为 CHAP 密钥。

规划客户机配置

当规划 iSCSI 客户机配置时, 您将需要以下信息:

- 要访问 SAN 的是哪些启动器 (及其 IQN) ?
- 如果计划使用 CHAP 验证, 则每个启动器都使用什么 CHAP 凭证?
- 需要多少 iSCSI 磁盘 (LUN) 及其容量?
- 是否需要在多个启动器之间共享 LUN?

为使设备能够使用 RADIUS 执行 CHAP 验证, 必须满足以下要求:

- 设备必须指定 RADIUS 服务器的地址, 以及与该 RADIUS 服务器通信时要使用的密钥
- RADIUS 服务器 (例如, 在其客户机文件中) 必须有一个条目指明此设备的地址并指定与上面相同的密钥
- RADIUS 服务器 (例如, 在其用户文件中) 必须有一个条目指明 CHAP 名称并与每个启动器的 CHAP 密钥匹配
- 如果启动器使用其 IQN 名称作为 CHAP 名称 (建议的配置), 则设备无需在每个 "Initiator" (启动器) 框中分别输入启动器条目 - RADIUS 服务器可执行所有验证步骤。
- 如果启动器使用单独的 CHAP 名称, 则设备必须有一个启动器条目对应于该启动器并指定从 IQN 名称到 CHAP 名称的映射关系。该启动器条目无需指定启动器的 CHAP 密钥。

iSCSI 故障排除

有关对常见 iSCSI 配置错误进行故障排除的技巧, 请参见[“iSCSI” \[189\]](#)一节。

监测 iSCSI 性能

可以通过 [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Analytics”监测 iSCSI 性能，在其中可按启动器、目标或 LUN 细分操作或吞吐量。

使用 BUI 配置 iSCSI

▼ 创建 Analytics 工作表

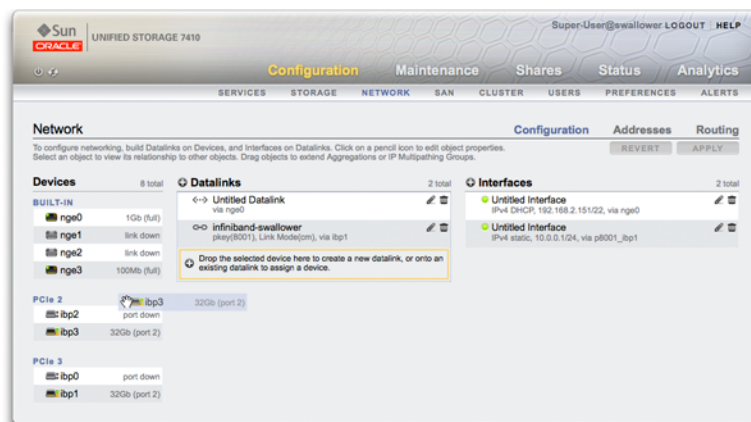
要创建 Analytics 工作表以按启动器监测操作，请执行以下操作：

1. 转至 "Analytics" 屏幕。
2. 单击 "Add Statistic"（添加统计信息）对应的  添加按钮。此时将显示包含所有统计信息的菜单。
3. 在该菜单的 "Protocols"（协议）部分下，选择 "iSCSI operations"（iSCSI 操作数） > "Broken down by initiator"（按启动器细分）。此时将显示按启动器细分的当前操作的图形。
4. 要查看更详细的 Analytics 信息，请从图形左侧的字段中选择相应启动器，然后单击  图标。此时将显示详细的 Analytics 菜单。

▼ iSER 目标配置

在 BUI 中，iSER 目标在 "Configuration"（配置） > "SAN" 屏幕上作为 iSCSI 目标进行管理。

1. 要配置 ibp(x) 接口，选择需要的 ibp(x) 接口（或 ipmp），并将其拖动到 "数据链路" 列表中以在 "Configuration"（配置） > "Network"（网络）屏幕上创建数据链路。
2. 将数据链路拖动到 "Interfaces"（接口）列表来创建新接口。

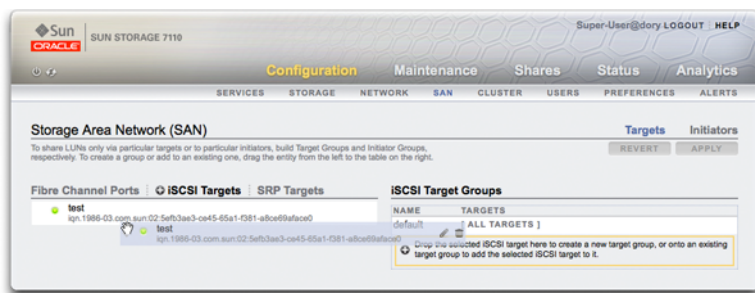


3. 要创建 iSER 目标，在 "Configuration" (配置) > "SAN" 页面上单击 "iSCSI Targets" (iSCSI 目标) 链路。
4. 要添加一个新的带别名的 iSER 目标，请单击  添加图标。
5. 要创建目标组，请将刚刚创建的目标拖动到 "iSCSI Target Group" (iSCSI 目标组) 列表。

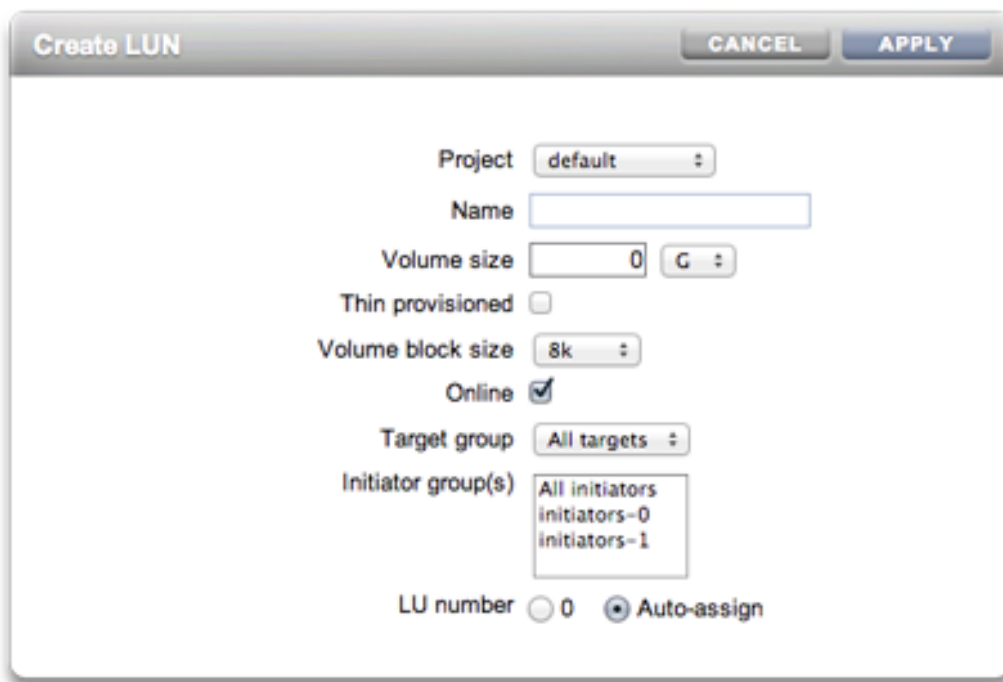


6. 要创建启动器，请单击 "Initiator" (启动器) 链接，然后单击 "iSCSI initiators" (iSCSI 启动器) 链接。
7. 要添加新启动器，单击  添加图标。
8. 输入启动器 IQN 和别名，然后单击 "OK" (确定)。尽管创建启动器组是可选操作，但是如果不创建组，与目标关联的 LUN 将可用于所有启动器。

9. 要创建组，请将启动器拖动到 "iSCSI Initiator Groups" (iSCSI 启动器组) 列表。



10. 要创建 LUN，在 "Shares" (共享资源) 页面上单击 LUN。
11. 单击  添加图标，并使用 "Target Group" (目标组) 和 "Initiator Group" (启动器组) 菜单将新 LUN 与已创建的目标或启动器组关联。
12. :



使用 CLI 配置 iSCSI

使用自动生成的 IQN 添加 iSCSI 目标

```

ahi:configuration san iscsi targets> create
ahi:configuration san iscsi targets target (uncommitted)> set alias="Target 0"
ahi:configuration san iscsi targets target (uncommitted)> set auth=none
ahi:configuration san iscsi targets target (uncommitted)> set interfaces=igbl
ahi:configuration san iscsi targets target (uncommitted)> commit
ahi:configuration san iscsi targets> list
TARGET      ALIAS
target-000 Target 0
|
+--> IQN
      iqn.1986-03.com.sun:02:daf0161f-9f5d-e01a-b5c5-e1efa9578416

```

添加使用特定 IQN 和 RADIUS 验证的 iSCSI 目标

```

ahi:configuration san iscsi targets> create
ahi:configuration san iscsi targets target (uncommitted)> set alias="Target 1"
ahi:configuration san iscsi targets target (uncommitted)>
    set iqn=iqn.2001-02.com.acme:12345
ahi:configuration san iscsi targets target (uncommitted)> set auth=radius
ahi:configuration san iscsi targets target (uncommitted)> set interfaces=igbl
ahi:configuration san iscsi targets target (uncommitted)> commit
ahi:configuration san iscsi targets> list
TARGET      ALIAS
target-000 Target 0
|
+--> IQN
      iqn.1986-03.com.sun:02:daf0161f-9f5d-e01a-b5c5-e1efa9578416
target-001 Target 1
|
+--> IQN
      iqn.2001-02.com.acme:12345

```

添加使用 CHAP 验证的 iSCSI 启动器

```

ahi:configuration san iscsi initiators> create
ahi:configuration san iscsi initiators initiator (uncommitted)>
    set initiator=iqn.2001-02.com.acme:initiator12345
ahi:configuration san iscsi initiators initiator (uncommitted)> set alias="Init 0"
ahi:configuration san iscsi initiators initiator (uncommitted)>
    set chapuser=thisismychapuser
ahi:configuration san iscsi initiators initiator (uncommitted)>
    set chapsecret=123456789012abc
ahi:configuration san iscsi initiators initiator (uncommitted)> commit
ahi:configuration san iscsi initiators> list

```

```
NAME          ALIAS
initiator-000 Init 0
|
+--> INITIATOR
      iqn.2001-02.com.acme:initiator12345
```

添加 iSCSI 目标组

```
ahi:configuration san iscsi targets groups> create
ahi:configuration san iscsi targets group (uncommitted)> set name=tg0
ahi:configuration san iscsi targets group (uncommitted)>
  set targets=iqn.2001-02.com.acme:12345,
    iqn.1986-03.com.sun:02:daf0161f-9f5d-e01a-b5c5-e1efa9578416
ahi:configuration san iscsi targets group (uncommitted)> commit
ahi:configuration san iscsi targets groups> list
GROUP      NAME
group-000  tg0
|
+--> TARGETS
      iqn.2001-02.com.acme:12345
      iqn.1986-03.com.sun:02:daf0161f-9f5d-e01a-b5c5-e1efa9578416
```

添加 iSCSI 启动器组

```
ahi:configuration san iscsi initiators groups> create
ahi:configuration san iscsi initiators group (uncommitted)> set name=ig0
ahi:configuration san iscsi initiators group (uncommitted)>
  set initiators=iqn.2001-02.com.acme:initiator12345
ahi:configuration san iscsi initiators group (uncommitted)> commit
ahi:configuration san iscsi initiators groups> list
GROUP      NAME
group-000  ig0
|
+--> INITIATORS
      iqn.2001-02.com.acme:initiator12345
```

SRP

SCSI RDMA 协议是受设备支持的协议，用于通过提供 RDMA 服务的网络（即 InfiniBand）共享基于 SCSI 的存储。

SRP 目标配置

SRP 端口与其他 IB 端口服务（例如 IPoIB 和 RDMA）共享。SRP 服务只能在目标模式下运行。SRP 目标具有以下可配置的属性。

表 6-5 SRP 目标属性

属性	说明
Target EUI (目标 EUI)	该目标的扩展唯一标识符 (Extended Unique Identifier, EUI)。EUI 由系统自动分配，等同于 SRP 端口服务运行所借助的 HCA GUID。
Alias (别名)	此目标的用户可读昵称。

除了这些属性以外，BUI 还指示目标是处于联机状态还是脱机状态：

表 6-6 SRP 目标状态图标

图标	说明
	目标处于联机状态
	目标处于脱机状态

群集注意事项

在群集平台上，对于高可用性（多路径）配置，应对等目标配置到同一目标组中。SRP 多路径 I/O 是启动器端配置选项。

启动器配置

SRP 启动器具有以下可配置的属性。

表 6-7 SRP 启动器属性

属性	说明
Initiator EUI (启动器 EUI)	此启动器的 EUI。
Alias (别名)	此启动器的用户可读昵称。

监测 SRP 性能

SRP 性能可以通过《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“Analytics”进行监测，在其中可按启动器或目标细分操作或吞吐量。[{{Server}}/wiki/images/cfg_san_srp.png](#)

使用 BUI 配置 SRP 目标

▼ SRP 目标配置

本过程介绍配置 SRP 目标的步骤。

1. 将 HCA 端口连接到 IB 接口。
2. : 目标由设备自动搜索。
3. 要创建目标组，请转至 "Configuration" (配置) > "SAN" 屏幕。
4. 单击 "Target" (目标) 链接，然后单击 SRP 目标。
5. : 此时将显示 "SRP targets" (SRP 目标) 页面。
6. 要创建目标组，请使用  移动图标将目标拖动到“目标组”列表。
7. 单击 "Apply" (应用) 。
8. (可选) 要在 "Initiator" (启动器) 屏幕上创建启动器和启动器组，请单击  图标，从启动器收集 GUID，为其分配名称，然后将其拖动到启动器组。
9. 要创建 LUN 并将其与前面步骤中创建的 SRP 目标和启动器关联，请转至 "Shares" (共享资源) 屏幕。
10. 单击 "LUN" 链接，然后单击 LUN  图标。使用 "Create LUN" (创建 LUN) 对话框中的 "Target Group" (目标组) 和 "Initiator Group" (启动器组) 菜单选择要与 LUN 关联的 SRP 组。

使用 CLI 配置 SRP 目标

以下示例说明如何使用 `configuration san targets srp groups` 上下文创建名为 `targetSRPgroup` 的 SRP 目标组：

```
swallower:configuration san targets srp groups> create
swallower:configuration san targets srp group (uncommitted)> set name=targetSRPgroup
name = targetSRPgroup (uncommitted)
swallower:configuration san targets srp group (uncommitted)>
set targets=eui.0002C903000489A4
targets = eui.0002C903000489A4 (uncommitted)
```

```

swallower:configuration san targets srp group (uncommitted)> commit
swallower:configuration san targets srp groups> list
GROUP      NAME
group-000  targetSRPgroup
|
+--> TARGETS
      eui.0002C903000489A4

```

以下示例说明如何使用 CLI shares CLI 上下文创建 LUN 并将其与 targetSRPgroup 关联：

```

swallower:shares default> lun mylun
swallower:shares default/mylun (uncommitted)> set targetgroup=targetSRPgroup
      targetgroup = targetSRPgroup (uncommitted)
swallower:shares default/mylun (uncommitted)> set volsize=10
      volsize = 10 (uncommitted)
swallower:shares default/mylun (uncommitted)> commit
swallower:shares default> list
Filesystems:
NAME      SIZE      MOUNTPPOINT
test      38K       /export/test
LUNs:
NAME      SIZE      GUID
mylun     10G       600144F0E9D19FFB00004B82DF490001

```


用户配置

本节介绍可管理设备的用户和管理授予用户的权限的角色，以及如何使用 BUI 或 CLI 将其添加到系统中。

用户可以是以下两者之一：

- 本地用户 – 其所有帐户信息都保存在设备中。
- 目录用户 – 使用现有 “NIS” [223] 或 “LDAP” [225] 帐户，将补充授权设置保存在设备中。这样便可授予现有 NIS 或 LDAP 用户登录和管理设备的特权。

尽管数据服务支持本地用户，但需牢记以下几个事项。

- 对于本地用户，您没有 UID 控制权。NFSv3 使用任何其他项目以及 NFSv4 使用 AUTH_SYS 会出现问题。
- 不支持本地组。
- 为数据用途定义本地用户还可以允许本地用户登录管理接口。

通过为用户分配定制角色来为其授予特权。

用户角色

角色是可分配给用户的一组特权。可能需要创建授权级别不同的管理员和操作员角色。可为员工分配满足其需求的任何角色，而不为其分配不必要的特权。

我们认为使用角色的安全性显著高于使用共享管理员密码，例如为每个人提供 *root* 用户密码。角色限制用户仅具有必要的授权，并且在《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“日志”日志中还会将其操作归于各自的用户名。

默认情况下，存在称为 “Basic administration”（基本管理）的角色，它包含非常基本的授权。

用户授权

通过授权，用户可执行特定任务，例如创建共享资源、重新引导设备和更新系统软件。授权分组成范围，每个范围可具有一组可选过滤器，用于缩小授权的范围。例如，可使用过滤器使某项授权只能重新启动 HTTP 服务，而不是重新启动所有服务。

下表显示了可用范围：

表 7-1 用户可用范围

范围 BUI	范围 CLI	授权示例	过滤器示例
Active Directory	ad	加入 Active Directory 域	域名
警报	alert	配置警报过滤器和阈值	.
Analytics	stat	显示深入分析，读取统计信息	深入分析
群集	cluster	将资源故障恢复到群集对等设备	.
数据集	dataset	管理 Analytics 数据集的各个方面	配置
硬件	hardware	使磁盘联机 and 脱机	
密钥库	KeyStore	配置密钥库。	.
联网	net	配置联网设备、数据链路和接口	.
项目和共享资源	nas	更改项目和共享资源的常规属性	池、项目、共享资源
角色	role	配置角色的授权	角色名
SAN	stmf	配置 SAN 的授权	
服务	svc	重新启动服务	服务名
共享资源属性模式	schema	修改属性模式	.
系统	appliance	重新引导设备	设备名称
更新	update	更新系统软件	.
用户	user	更改密码	用户名
工作流程	workflow	修改工作流程	工作流程名称
工作表	worksheet	修改工作表	工作表名称

在 BUI 中浏览范围可查看存在的其他授权。目前提供了五十多种不同的授权，将来设备软件更新时可能会添加更多授权。

管理用户属性

管理用户和角色时可设置以下属性。

用户属性

添加用户时可设置以下所有属性，编辑用户时可设置其中部分属性：

表 7-2 用户属性

属性	说明
Type (类型)	"Directory" (目录) (从 NIS 或 LDAP 访问凭证) 或 "Local" (本地) (将用户保存在该设备中)
Username (用户名)	用户的唯一名称
Full Name (全名)	用户说明
Password/Confirm (密码/确认)	对于本地用户，请在这两个字段中键入初始密码
Require session annotation (需要会话注释)	启用后，用户登录设备时，必须提供登录目的文本说明。该注释可用于跟踪针对票证处理系统中的请求执行的工作，票证 ID 可用作会话注释。会话注释显示在《 Oracle ZFS Storage Appliance 客户服务手册 》中的“日志”日志中。
Kiosk user (Kiosk 用户)	启用后，用户只能查看 "Kiosk screen" (Kiosk 屏幕) 设置中的屏幕。例如，该属性可用于仅限用户查看“ 显示板 ” [46]。Kiosk 用户将无法通过 CLI 访问设备。
Kiosk screen (Kiosk 屏幕)	启用 "Kiosk user" (Kiosk 用户) 后，仅限该 Kiosk 用户查看的屏幕
Roles (角色)	该用户具有的角色
Exceptions (例外)	这些授权不属于通常因选定角色而具有的授权

角色属性

管理角色时可设置以下属性：

表 7-3 角色属性

属性	说明
Name (名称)	显示在列表中的角色名称
Description (说明)	角色的详细说明 (如果需要)

属性	说明
Authorizations (授权)	该角色的授权

用户 BUI 页面

"BUI Users" (BUI 用户) 页面将列出用户和组，以及用于管理的按钮。将鼠标悬停在条目上方可显示其克隆、编辑和销毁按钮。双击条目可查看其编辑屏幕。按钮如下所示：

表 7-4 用户 BUI 页面图标

图标	说明
	添加新用户/角色。该图标将显示新对话框，可在其中输入所需属性。
	显示搜索框。输入搜索字符串并按 Enter 键可在用户/角色列表中搜索该文本，并且仅显示匹配的条目。再次单击该图标或 "Show All" (全部显示) 可返回完整列表。
	克隆用户/角色。添加新用户/角色，起始字段基于该条目中的值
	编辑用户/角色
	删除用户/角色/授权

使用 BUI 配置用户

▼ 添加管理员

1. 检查相应的管理员角色是否列在 "Roles" (角色) 列表中。如果未列出，请添加角色 (请参见单独的任务)。
2. 单击 "Users" (用户) 旁边的  添加图标。
3. 设置用户属性。
4. 单击管理员角色对应的复选框。
5. 单击对话框顶部的 "Add" (添加) 按钮。新用户即会显示在 "Users" (用户) 列表中。

▼ 添加角色

1. 单击 "Roles" (角色) 旁边的  添加图标。
2. 设置角色的名称和说明。
3. 为角色添加授权 (请参见单独的任务)。
4. 单击对话框顶部的 "Add" (添加) 按钮。新角色即会显示在 "Roles" (角色) 列表中。

▼ 为角色添加授权

1. 选择 "Scope" (范围)。如果存在该范围的过滤器，它们将显示在 "Scope" (范围) 选择器下方。
2. 选择过滤器 (如果适用)。
3. 单击要添加的所有授权对应的复选框。
4. 单击 "Authorization" (授权) 部分中的 "Add" (添加) 按钮。授权即会添加到对话框的底部列表中。

▼ 从角色中删除授权

1. 将鼠标悬停在 "Roles" (角色) 列表中角色的上方，并单击  编辑图标。
2. 将鼠标悬停在底部列表中授权的上方，然后单击右侧的  垃圾箱图标。
3. 单击对话框顶部的 "Apply" (应用) 按钮。

▼ 添加只能查看显示板的用户

1. 添加目录用户或本地用户 (请参见单独的任务)。
2. 将 Kiosk 模式设置为 "true"，并检查是否已将 Kiosk 屏幕设置为 "status/dashboard" (状态/显示板)。

3. 用户现在应该可以登录，但是只能查看显示板。

使用 CLI 配置用户

可在 BUI 中执行的操作也可在 CLI 中执行。浏览用户、角色和授权管理时，键入 help 可列出可用的命令。

CLI 用户配置示例

要显示 CLI 用户和角色界面，可按照以下示例将 NIS 用户 "brendan" 添加到系统中，并授予重新启动 HTTP 服务的权限。其中包括创建该授权对应的角色。

首先创建称为 "webadmin" 的角色：

```
caji:> configuration roles
caji:configuration roles> role webadmin
caji:configuration roles webadmin (uncommitted)> set
      description="web server administrator"
              description = web server administrator (uncommitted)
caji:configuration roles webadmin (uncommitted)> commit
caji:configuration roles> show
Roles:

NAME            DESCRIPTION
basic           Basic administration
webadmin        web server administrator
```

我们已创建 "webadmin" 角色，现在将添加重新启动 HTTP 服务的授权。本示例还显示了 Tab 补齐功能的输出，其中列出了有效的输入，这在确定有效的范围和过滤器选项时非常有用：

```
caji:configuration roles> select webadmin
caji:configuration roles webadmin> authorizations
caji:configuration roles webadmin authorizations> create
caji:configuration roles webadmin auth (uncommitted)> set scope=tab
ad      cluster  net      schema  update
alert   hardware  replication  stat    user
appliance  nas      role     svc      worksheet
caji:configuration roles webadmin auth (uncommitted)> set scope=svc
      scope = svc
caji:configuration roles webadmin auth (uncommitted)> show
Properties:
      scope = svc
      service = *
allow_administer = false
allow_configure = false
allow_restart = false

caji:configuration roles webadmin auth (uncommitted)> set service=tab
```

```

*          ftp          ipmp          nis          ssh
ad          http          iscsi          ntp          tags
smb          identity      ldap          routing      vscan
datalink:igb0 idmap          ndmp          scrk
dns          interface:igb0 nfs          snmp
caji:configuration roles webadmin auth (uncommitted)> set service=http
                        service = http (uncommitted)
caji:configuration roles webadmin auth (uncommitted)> set allow_restart=true
                        allow_restart = true (uncommitted)
caji:configuration roles webadmin auth (uncommitted)> commit
caji:configuration roles webadmin authorizations> list
NAME      OBJECT                                PERMISSIONS
auth-000   svc.http                                restart

```

现在即已创建角色，接下来可进入用户部分来创建用户 "brendan" 并分配角色 "webadmin"：

```

caji:configuration roles webadmin authorizations> cd ../../..
caji:configuration> users
caji:configuration users> netuser brendan
caji:configuration users> show
Users:

NAME                USERNAME          UID      TYPE
Brendan Gregg       brendan           130948   Dir
Super-User           root              0        Loc

caji:configuration users> select brendan
caji:configuration users brendan> show
Properties:
        logname = brendan
        fullname = Brendan Gregg
        initial_password = *****
        require_annotation = false
        roles = basic
        kiosk_mode = false
        kiosk_screen = status/dashboard

Children:
        exceptions => Configure this user's exceptions
        preferences => Configure user preferences
caji:configuration users brendan> set roles=basic,webadmin
        roles = basic,webadmin (uncommitted)
caji:configuration users brendan> commit

```

用户 "brendan" 现在应该可以使用其 NIS 密码登录，并可在设备上重新启动 HTTP 服务。

▼ 添加管理员

1. 转至 **configuration roles**。

2. 键入 `show`。通过针对每个角色运行 `select` 然后运行 `authorizations show`，查找具有相应管理授权的角色。如果不存在相应的角色，请首先创建该角色（请参见单独的任务）。
3. 转至 `configuration users`。
4. 对于目录用户 (NIS、LDAP)，键入 `netuser`，后跟要添加的现有用户名。对于本地用户，键入 `user`，后跟要添加的用户名；然后键入 `show` 查看需要设置的属性。键入 `set`，然后键入 `commit`。
5. 此时即已创建用户，但是尚未定制其所有属性。键入 `select`，后跟其用户名。
6. 接下来键入 `show` 查看首选项的完整列表。现在即可添加角色和授权例外以及[第 8 章 设置 ZFSSA 首选项](#)。

▼ 添加角色

1. 转至 `configuration roles`。
2. 键入 `role`，后跟要创建的角色名称。
3. 设置说明，然后键入 `commit` 提交角色。
4. 为角色添加授权（请参见单独的任务）。

▼ 为角色添加授权

1. 转至 `configuration roles`。
2. 键入 `select`，后跟角色名称。
3. 键入 `authorizations`。
4. 键入 `create` 添加授权
5. 键入 `set scope=`，后跟范围名称。使用 Tab 补齐功能查看列表。
6. 键入 `show` 查看可用的过滤器和授权。

7. 键入 `set` 将所需授权设置为 "true"，然后设置过滤器（如果有）。Tab 补齐功能可帮助显示有效的过滤器设置。
8. 键入 `commit`。现在即已添加授权。

▼ 从角色中删除授权

1. 转至 `configuration roles`。
2. 键入 `select`，后跟角色名称。
3. 键入 `authorizations`。
4. 键入 `show` 列出授权。
5. 键入 `destroy`，后跟授权名称（例如 "auth-001"）。现在即已销毁授权。

设置 ZFSSA 首选项

本节包含语言环境、会话属性和 SSH 密钥的首选项设置。

首选项属性

登录 BUI 后，可为您的帐户设置以下首选项，但是无法设置其他用户帐户首选项。

表 8-1 首选项设置

属性	说明
Initial login screen（初始登录屏幕）	成功登录后 BUI 将加载的第一个页面。默认情况下，这是“ 状况显示板 ” [46]。
Locality（语言环境）	默认情况下为 C。C 和 POSIX 语言环境仅支持 ASCII 字符或纯文本。ISO 8859-1 支持以下语言：南非荷兰语、巴斯克语、加泰罗尼亚语、丹麦语、荷兰语、英语、法罗语、芬兰语、法语、加利西亚语、德语、冰岛语、爱尔兰语、意大利语、挪威语、葡萄牙语、西班牙语和瑞典语。
Session timeout（会话超时）	离开 BUI 后浏览器将自动注销会话的时间
Current session annotation（当前会话注释）	添加到审计日志中的注释文本
Advanced analytics statistics（高级 Analytics 统计信息）	这样提供的统计信息将比《 Oracle ZFS Storage Appliance Analytics 指南 》中的“Analytics”中多
SSH Public Keys（SSH 公钥）	RSA/DSA 公钥。文本注释可与密钥关联，以帮助管理员跟踪添加原因。在 BUI 中，这些密钥仅适用于当前用户，要为用户添加密钥，请使用 CLI。

使用 CLI 设置首选项

可在 CLI 中的 `configuration users` 下设置首选项。以下示例显示了如何为 “brendan” 用户帐户启用高级 Analytics：

```
caji:> configuration users
caji:configuration users> select brendan
caji:configuration users brendan> preferences
caji:configuration users brendan preferences> show
Properties:
    locale = C
    login_screen = status/dashboard
    session_timeout = 15
    advanced_analytics = false

Children:
    keys => Manage SSH public keys

caji:configuration users brendan preferences> set advanced_analytics=true
    advanced_analytics = true (uncommitted)
caji:configuration users brendan preferences> commit
```

可在 CLI 中的 `configuration preferences` 下设置您自己的首选项。以下示例显示了如何为您自己的帐户设置会话注释：

```
twofish:> configuration preferences
twofish:configuration preferences> show
Properties:
    locale = C
    login_screen = status/dashboard
    session_timeout = 15
    session_annotation =
    advanced_analytics = false

Children:
    keys => Manage SSH public keys

twofish:configuration preferences> set session_annotation="Editing my user preferences"
    session_annotation = Editing my user preferences (uncommitted)
twofish:configuration preferences> commit
```

使用 CLI 设置 SSH 公钥

从其他主机自动执行 CLI 脚本时可能需要 SSH 公钥。以下示例显示了如何从 CLI 添加 SSH 密钥：

```
caji:> configuration preferences keys
caji:configuration preferences keys> create
caji:configuration preferences key (uncommitted)> set type=DSA
caji:configuration preferences key (uncommitted)> set key="...DSA key text..."
    key = ...DSA key text...== (uncommitted)
caji:configuration preferences key (uncommitted)> set comment="fw-log1"
    comment = fw-log1 (uncommitted)
caji:configuration preferences key (uncommitted)> commit
caji:configuration preferences keys> show
Keys:
```

NAME	MODIFIED	TYPE	COMMENT
key-000	10/12/2009 10:54:58	DSA	fw-log1

密钥文本仅指密钥文本自身（通常是数百个字符），不含空格。

警报配置

本节介绍系统警报、如何对其进行定制，以及在何处可找到警报日志。要通过《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“Analytics”监视统计信息，请创建定制阈值警报。要配置系统以响应某些类型的警报，请使用警报操作。

警报类别

重要的设备事件可触发警报，其中包括硬件和软件故障。这些警报包括在《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“日志”中，也可以配置用于执行任意警报操作。

警报分为以下类别：

类别	说明
群集	群集事件，包括链路故障和对等错误
定制	从定制警报配置生成的事件
硬件事件	设备引导和硬件配置更改
硬件故障	任何硬件故障
NDMP 操作	备份和恢复开始与完成的事件。该组以 "NDMP: backup only" 和 "NDMP: restore only" 的形式提供，仅用于备份或恢复事件
网络	网络端口、数据链路和 IP 接口事件与故障
回拨	支持包上载事件
远程复制	发送和接收事件与故障。该组以 "Remote replication: source only" 和 "Remote replication: target only" 的形式提供，仅用于源事件或目标事件
服务故障	软件第 11 章 ZFSSA 服务故障事件
阈值	基于《 Oracle ZFS Storage Appliance Analytics 指南 》中的“Analytics”统计信息的定制警报

类别	说明
ZFS 池	存储池事件，包括清理和热空间激活

支持的警报操作

支持以下操作。

发送电子邮件

可发送包含警报详细信息的电子邮件。配置需要电子邮件地址和电子邮件主题行。以下是根据阈值警报发送的电子邮件样例：

```
From aknobody@caji.com Mon Oct 13 15:24:47 2009
Date: Mon, 13 Oct 2009 15:24:21 +0000 (GMT)
From: Appliance on caji <noreply@caji.com>
Subject: High CPU on caji
To: admin@hostname.com

SUNW-MSG-ID: AK-8000-TT, TYPE: Alert, VER: 1, SEVERITY: Minor
EVENT-TIME: Mon Oct 13 15:24:12 2009
PLATFORM: i86pc, CSN: 0809QAU005, HOSTNAME: caji
SOURCE: svc:/appliance/kit/akd:default, REV: 1.0
EVENT-ID: 15a53214-c4e7-eae4-dae6-a652a51ea29b
DESC: cpu.utilization threshold of 90 is violated.
AUTO-RESPONSE: None.
IMPACT: The impact depends on what statistic is being monitored.
REC-ACTION: The suggested action depends on what statistic is being monitored.

SEE: https://192.168.2.80:215/#maintenance/alert=15a53214-c4e7-eae4-dae6-a652a51ea29b
```

有关设备如何发送邮件的详细信息可在“SMTP” [250]服务屏幕上进行配置。

发送 SNMP 陷阱

如果在“SNMP” [251] 服务中配置了 SNMP 陷阱目标且该服务处于联机状态，则可发送包含警报详细信息的 SNMP 陷阱。以下是通过 Net-SNMP 工具 snmptrapd -P 看到的 SNMP 陷阱示例：

```
# /usr/sfw/sbin/snmptrapd -P
2009-10-13 15:31:15 NET-SNMP version 5.0.9 Started.
2009-10-13 15:31:34 caji.com [192.168.2.80]:
    iso.3.6.1.2.1.1.3.0 = Timeticks: (2132104431) 246 days, 18:30:44.31
    iso.3.6.1.6.3.1.1.4.1.0 = OID: iso.3.6.1.4.1.42.2.225.1.3.0.1
    iso.3.6.1.4.1.42.2.225.1.2.1.2.36.55.99.102.48.97.99.100.52.45.51.48.
99.49.45.52.99.49.57.45.101.57.99.98.45.97.99.50.55.102.55.49.50.54.
98.55.57 = STRING: "7cf0acd4-30c1-4c19-e9cb-ac27f7126b79"
```

```
iso.3.6.1.4.1.42.2.225.1.2.1.3.36.55.99.102.48.97.99.100.52.45.51.48.
99.49.45.52.99.49.57.45.101.57.99.98.45.97.99.50.55.102.55.49.50.54.
98.55.57 = STRING: "alert.ak.xmlrpc.threshold.violated"
iso.3.6.1.4.1.42.2.225.1.2.1.4.36.55.99.102.48.97.99.100.52.45.51.
48.99.49.45.52.99.49.57.45.101.57.99.98.45.97.99.50.55.102.55.49.50.
54.98.55.57 = STRING: "cpu.utilization threshold of 90 is violated."
```

发送 Syslog 消息

如果启用了 syslog 服务，则可将包含警报详细信息的 syslog 消息发送到一个或多个远程系统。有关 syslog 有效载荷示例和如何在其他操作系统上配置 syslog 接收方的说明，请参阅介绍“[Syslog 转发服务](#)”[254]的文档。

恢复/暂停数据集

可以恢复或暂停 Analytics [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“数据集”。这在跟踪偶发性性能问题且全天候启用这些数据集并不可取时特别有用。

例如：假设您注意到 CPU 活动每周出现一次或两次高峰，并且其他 Analytics 信息表明由此造成了 NFS 性能下降。您可启用某些其他数据集，但是没有十分足够的信息证明出了什么问题。如果可按主机名和文件名数据集启用 NFS，必然能够更好地了解根源。但是，这些特殊的数据集可能很难处理 – 全天候启用会导致全面的性能下降。这正是恢复/暂停数据集操作的用武之地。可配置一个阈值警报，仅在检测到 CPU 活动高峰时按主机名和文件名数据集恢复暂停的 NFS；再配置第二个警报，随后在收集短时间间隔内的数据后暂停这些数据集。最终结果 – 仅在问题出现时收集所需数据，最大限度地降低了这些数据收集对性能的影响。

恢复/暂停工作表

这些操作用于恢复或暂停整个 Analytics [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“[Open Worksheets \(打开的工作表\)](#)”，其中可包含多个数据集。这样做的原因与恢复和暂停数据集的原因类似。

执行工作流

可选择性地执行工作流作为警报操作。要使工作流符合作为警报操作的资格，其 alert 操作必须设置为 true。有关详细信息，请参阅“[工作流作为警报操作](#)”[377]。

阈值警报

这些警报基于来自 [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Analytics”的统计信息。以下是创建阈值警报时的属性：

表 9-2 阈值警报属性

属性	描述
Threshold (阈值)	阈值统计信息来自《Oracle ZFS Storage Appliance Analytics 指南》中的“Analytics”，是自描述性信息（例如 "Protocol: NFSv4 operations per second"）
exceeds/falls below (超过/低于)	定义阈值与当前统计信息相比如何
Timing: for at least (计时：至少)	当前统计值必须超过/低于阈值的持续时间
only between/only during (仅之间/仅过程中)	可设置这些属性以便仅在一天当中的某些时间（例如工作时间）发送阈值
Repost alert every ... this condition persists. (在此条件持续存在的情况下每隔 ... 重新发布警报。)	启用后，存在突破阈值的情况时，每过设定的一段时间间隔会重新执行一次警报操作（例如发送电子邮件）
Also post alert when this condition clears for at least ... (此状态清除之后距再次发布警报的最少时间 ...)	突破阈值的情况消除至少设定的一段时间间隔后发送跟进警报

"Add Threshold Alert" (添加阈值警报) 对话框已经过组织，使其好像描述警报的段落一样方便读取。默认读取内容：

"Threshold CPU: percent utilization exceeds 95 percent"

"Timing for at least 5 minutes only between 0:00 and 0:00 only during weekdays"

"Repost alert every 5 minutes while this condition persists."

"Also post alert when this condition clears for at least 5 minutes"

使用 BUI 配置警报

"Configuration" (配置) > "Alerts" (警报) 页面顶部是 "Alert Actions" (警报操作) 和 "Threshold Alerts" (阈值警报) 对应的选项卡。有关在 BUI 中配置这些内容的逐步说明，请参见“任务”。

▼ 添加阈值警报

1. 单击 "Threshold alerts" (阈值警报) 旁边的添加图标。
2. 选择要监视的统计信息。可以使用《Oracle ZFS Storage Appliance Analytics 指南》中的“统计信息”查看统计信息，以检查其是否合适。
3. 选择 "exceeds/falls below" (超出/低于) 以及所需值。

4. 输入 "Timing" (计时) 详细信息。默认情况下, 仅在突破阈值至少 5 分钟时发布警报, 每 5 分钟重新发布一次, 并在突破阈值的情况消除 5 分钟后发布警报。
5. 从下拉菜单中选择警报操作, 然后填写右侧的必填字段。
6. 如果需要, 单击 "Alert actions" (警报操作) 旁边的添加图标继续添加警报操作。
7. 单击对话框顶部的 "APPLY" (应用)。

▼ 添加警报操作

1. 单击 "Alert actions" (警报操作) 旁边的添加图标。
2. 选择 "Category" (类别), 或针对所有内容选择 "All events" (所有事件)。
3. 选择 "All Events" (所有事件) 或 "a Subset of Events" (事件子集)。如果选择了子集, 请定制复选框列表, 使其与所需警报事件匹配。
4. 使用 "Alert actions" (警报操作) 中的下拉菜单选择警报类型。
5. 输入警报操作的详细信息。单击 "TEST" (测试) 按钮可创建测试警报并执行该警报操作 (在检查电子邮件或 SNMP 是否配置正确时非常有用)。
6. 单击 "Alert actions" (警报操作) 旁边的添加图标可添加多项警报操作。
7. 单击右上方的 "ADD" (添加)。

使用 CLI 配置警报

也可使用 `configuration alerts` 上下文从 CLI 配置警报。有关在 CLI 中配置这些内容的逐步说明, 请参见“任务”。

▼ 添加阈值警报

1. 输入 `configuration alerts thresholds` 上下文, 并输入 `create` 命令。
2. 输入 `"set statname=[name]"`, 其中 [name] 是要监视的所需统计信息。要确定 CLI 名称, 输入 `"set statname="`, 然后按 Tab 键。有关各项统计信息的详细信息, 请参见《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“统计信息”并单击统计信息名称。

3. 输入 `"set limit=[number]"`，其中 `[number]` 是所需的阈值。
4. 输入 `commit`。如果要稍后为此阈值警报添加警报操作，请记录“沙漏”标识符和阈值 ID。
5. 输入 `list` 以确定新阈值警报的名称，包括其编号。查找具有您刚设置的限制和统计信息名称的阈值。
6. 输入 `"select threshold-[number]"`，其中 `[number]` 是上一步中标识的同一数值。
7. 输入 `list`。如有必要，立即更正参数。默认情况下，最小的发布参数、频率参数和最小清除参数都设置为 5 分钟。这意味着仅在突破阈值至少 5 分钟时发布警报，每 5 分钟重新发布一次，并在突破阈值的情况消除 5 分钟后发布警报。
8. 输入 `done`，然后再次输入 `done`。

▼ 添加警报操作

1. 输入 `configuration alerts actions` 上下文，并输入 `create` 命令。
2. 输入 `get category = (unset)` 命令以转至 "category" (类别) 属性。
3. 输入 `set category=thresholds`。
4. 输入 `"set thresholdid=[id]"`，其中 `[id]` 是为阈值警报自动创建的标识符。
5. 输入 `commit`。
6. 输入 `list` 以确定新警报操作的名称，包括其编号。查找无已分配操作和处理程序的阈值。
7. 输入 `"select actions-[number]"`，其中 `[number]` 是上一步中标识的同一数值。
8. 输入 `action`，然后输入 `get`。
9. 默认情况下，警报类型为电子邮件。如果接受默认设置，请跳到下一个步骤。否则，请输入 `"set handler=[type]"`，其中 `[type]` 为 `snmptrap`、`syslog`、`resumedataset`、`suspenddataset`、`resumeworksheet`、`suspendworksheet` 或 `executeworkflow`。然后输入 `get` 查看所需参数。仅 `snmptrap` 和 `syslog` 没有参数。
10. 设置各个所需参数。例如，要设置电子邮件警报主题行，可输入 `"set subject=[subject]"`，其中 `[subject]` 是所需的电子邮件主题行。
11. 使用 `show` 命令以确保所有参数都已输入。

12. 输入 `commit`，然后输入 `list`。如有必要，立即更正参数。
13. 输入 `done`，然后再次输入 `done`。

群集配置

Oracle ZFS Storage Appliance 支持设备的协作群集。该策略可以是可用性增强综合方法的一部分，这种方法还可能包括内置到所有 Oracle ZFS Storage Appliance 中的客户端负载平衡、妥善的场地规划、主动性和响应式维护与修复以及单设备硬件冗余。

群集功能依赖于对存储资源的共享访问。要配置群集，两个机头必须是同一型号。请注意，7420（带有 2GHz 或 2.40GHz CPU）基于同一平台，可与 7420（带有 1.86GHz 或 2.00GHz CPU）形成群集。

群集的功能与优势

了解 Oracle ZFS Storage Appliance 群集实施的范围非常重要。“群集”一词在业内用于指用途多种多样的多项不同技术。本文中的群集指这样一种元系统：由两个设备机头和共享存储组成，用于在其中一个机头发生某些硬件或软件故障的情况下提高可用性。一个群集只包含两个设备或存储控制器，为简洁起见，本文档中通篇将其称为机头。可以从群集可用的资源集中为每个机头分配一组存储、网络和其他资源，以便构建两个主要拓扑之一。许多人使用主动-主动一词描述这样一种群集：包含两个（或更多）存储池，为每个机头分配其中一个存储池以及客户机用于连接到该池中所存储数据的网络资源。而主动-被动是指将单个存储池以及关联的网络接口分配给指定为主动的机头。这两个拓扑都受 Oracle ZFS Storage Appliance 支持。两者之间的区别是人为的，没有软件或硬件差别，只需添加或销毁存储池即可随意转换。在这两种情况下，如果一个机头发生故障，另一个（其对应设备）将控制所有已知资源，并提供与这些资源关联的服务。

通过群集，对等设备可在执行修复或更换时提供服务，而不会在修复机头时停机数小时甚至数天时间。此外，群集还支持软件的滚动升级，这样可减少迁移到较新的软件时对业务的中断。除可用性增强之外，一些群集技术还具备某些其他功能。Oracle ZFS Storage Appliance 群集子系统目的并不在于提供这些功能。特别需要指出的是，它不具备以下功能：在多个机头之间实现负载平衡、在发生存储故障时提高可用性、为客户机提供多个设备之间统一的文件系统名称空间，或为了灾难恢复目的而在广泛的地理区域内划分服务职责。这些功能同样不在本文档论述范围之内，但是 Oracle ZFS Storage Appliance 和它提供的数据协议支持可提高可用性的多种其他功能和策略：

- 数据的[第 13 章 复制](#)，可用于一个或多个远程站点（从地理位置上讲）的灾难恢复
- 数据的客户端镜像，可使用多个任意位置的存储服务器提供的冗余“iSCSI”[\[189\]](#) LUN 执行

- 负载平衡，内置到“NFS” [185] 协议中，可由外部硬件或软件为其他某些协议提供（仅适用于只读数据）
- 冗余硬件组件，包括电源、网络设备和存储控制器
- 《Oracle ZFS Storage Appliance 客户服务手册》中的“问题”软件，可识别发生故障的组件、将其从服务中删除，以及指导技术人员修复或更换适当的硬件
- 由 LACP 和“IPMP” [242] 功能提供的网络结构冗余
- 冗余存储设备 (RAID)

有关其他可用性功能的其他信息可在本文档相应章节中找到。

群集的缺点

在群集与单机 Oracle ZFS Storage Appliance 配置之间选择时，权衡群集操作的利与弊非常重要。整个 IT 行业的常见做法是将群集视为自动化架构决策，但这种想法反映了该领域内某些供应商推崇的群集风险和回报理想化的观点。除与第二个机头关联的显著提高的前期和持续硬件与支持成本之外，群集还具有其他技术和操作风险。其中某些风险可通过确保对所有人员进行有关群集操作的全面培训来减轻，而其他一些则是群集操作概念固有的风险。此类风险包括：

- 接管期间应用程序可能无法容忍依赖于协议的行为
- 群集软件自身可能会发生故障或导致另一个子系统发生故障，这种情况在单机操作中不会发生
- 提高管理复杂性以及执行管理任务时操作员出错的可能性
- 多种故障或严重的操作员错误可能会导致数据丢失或损坏，这种情况在单机配置中不会发生
- 增加从意外的软件和/或硬件状态恢复的难度

这些是基本的成本和风险，以某种形式存在于市场上所有群集产品或支持群集的产品（包括 Oracle ZFS Storage Appliance）中，无法减轻或完全消除。存储架构师必须将其与群集的以下主要优势进行权衡：发生相当少见的灾难性硬件或软件故障事件时，可将不可用期限从数小时或数天缩短到几分钟甚至更短时间。这种成本/优势分析是否表明在 Oracle ZFS Storage Appliance 部署中使用群集更有利取决于一些局部因素，例如 SLA 条款、现有支持人员及其资格、预算限制、认为发生各种可能故障的可能性，以及增强可用性的备选策略的适用性。这些因素高度依赖于站点、应用和业务，必须逐个进行评估。了解本节中其余部分的内容有助于您在设计和实施统一存储基础结构期间做出适当的选择。

群集术语

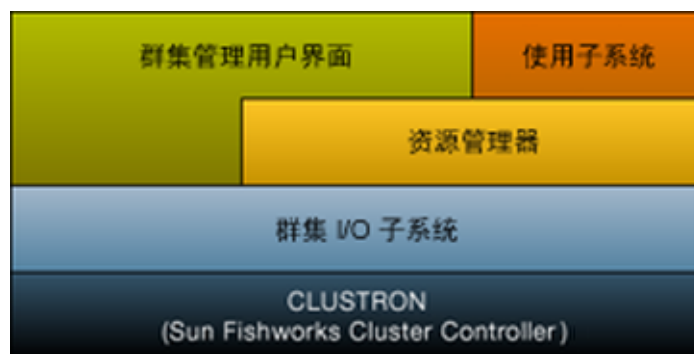
本文档通篇使用此处定义的术语。在大多数情况下，将提供更多的上下文和详细信息以及更广泛的概念对其进行解释。群集状态和资源类型将在下一节进行介绍。请根据需要返回本节加以参考。

- 导出：使特定机头上的资源处于非活动状态的过程
- 故障恢复：从 AKCS_OWNER 状态转为 AKCS_CLUSTERED 的过程，期间所有外部资源（分配给对等设备的资源）都将导出，然后由对等设备导入
- 导入：使特定机头上的资源处于活动状态的过程
- 对等设备：群集中的其他设备
- 重新链接：从对等设备中检索并重新同步资源映射
- 资源：存在于一个或两个机头上且可能处于活动状态的物理或虚拟对象
- 接管：从 AKCS_CLUSTERED 或 AKCS_STRIPPED 状态转为 AKCS_OWNER 的过程，期间导入所有资源

了解群集

并入系列的群集子系统由三个主要构建块组成（参见图 1）。群集 I/O 子系统和硬件设备为群集内机头间的通信提供传输，并负责监视对等设备的状态。这种传输由资源管理器使用，便于数据服务提供程序和其他管理子系统与群集系统连接。最后，群集管理用户界面提供设置任务、资源分配和指定、监视以及接管和故障恢复操作。下面几节将对每个构建块进行详细介绍。

图 10-1 群集子系统



群集互连 I/O

所有机头间通信都是通过 CLUSTERON 硬件提供的三条群集 I/O 链路中的一条链路发送一条或多条消息（参见下图）。该设备提供两条低速串行链路和一条以太网链路。使用串行链路可提高可靠性；系统负荷极其繁重时，可能无法足够快速地为以太网链路提供

服务。要使群集系统响应负荷，执行误报故障检测和不需要的接管效果最差。接管期间不会处理请求，而是由客户机将其加入队列，导致接管后除已经非常繁重的负荷之外，还会存在大量延迟的请求。Oracle ZFS Storage Appliance 使用的串行链路不易受这种故障模式的影响。以太网链路可以为非心跳消息提供性能较高的传输（例如重新链接同步），并提供备份心跳。

三条链路全部使用普通的直通 EIA/TIA-568B（8 线，千兆位以太网）电缆构成。为了能够在两个完全相同的控制器之间使用直通电缆，必须使用电缆连接两个连接器上相反的插槽，如下面有关布线的一节中所示。

图 10-2 ZS3-2 控制器群集 I/O 端口



表 10-1 ZS3-2 控制器群集 I/O 端口

图例			
1 串行端口 0	2 串行活动 LED 指示灯	3 串行状态 LED 指示灯	4 以太网端口
5 串行端口 1	6 以太网状态 LED 指示灯	7 以太网活动 LED 指示灯	

图 10-3 ZS3-4 和 7x20 控制器群集 I/O 端口

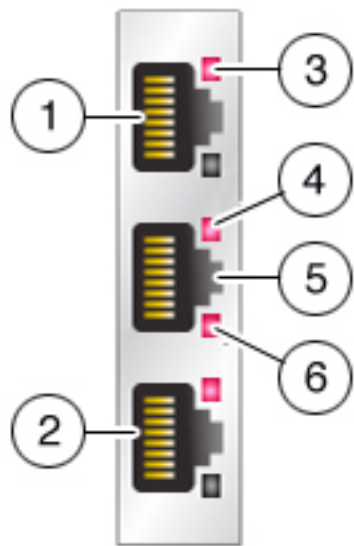


图 2.ZS3-4 和 7x20 控制器群集 I/O 端口

表 10-2 ZS3-4 和 7x20 控制器群集 I/O 端口

图例			
1 串行端口 1	2 串行端口 0	3 串行状态 LED 指示灯	4 以太网状态 LED 指示灯
5 以太网端口	6 以太网活动 LED 指示 灯		

群集机头只通过由群集互连建立的安全专用网络互相通信，从不通过打算用于服务或管理的网络接口进行通信。消息归为两大类：一类是用于检测远程机头故障的定期心跳，另一类是与资源管理器和群集管理子系统关联的较高级别的流量。心跳要在全部三条链路上发送，它们以固定的时间间隔持续传送，从不对其进行确认或重新传送，因为所有心跳均完全相同而不包含任何唯一的信息。其他流量可通过任何链路发送，通常是传送时可用的速度最快的链路，并且为使较高级别的软件保持可靠的传输，要对这些流量进行确认、验证和重新传送。

无论类型或来源为何，每条消息都作为单个 128 字节的数据包发送，并包含 1 到 68 字节的数据有效载荷和 20 字节的验证散列，以确保数据完整性。串行链路以 115200 bps

的速率运行（包含 9 个数据位以及 1 个起始位和停止位）；以太网链路以 1 Gbps 的速率运行。因此，串行链路上的有效消息延迟约为 12.2 毫秒。以太网延迟差别很大，通常延迟约为几微秒，但设备管理软件上的有效延迟可能会因系统负荷而高得多。

通常，每个机头都在全部三条群集 I/O 链路上以 50 毫秒的时间间隔发送心跳消息。如果 200 毫秒（串行链路）或 500 毫秒（以太网链路）后未能收到任何消息，将视为发生了链路故障。如果全部三条链路都发生了故障，则认为对等设备发生了故障，将执行接管仲裁。如果发生紧急情况，紧急状况下的机头将通过每条串行链路传送一条通知消息；无论其他任何链路处于哪种状态，对等设备都会立即开始接管。考虑到这些特征，群集子系统通常可以在以下时间内检测到对等设备发生故障：

- 550 毫秒，如果对等设备已停止响应或断电，或者
- 30 毫秒，如果对等设备遇到致命的软件错误而导致操作系统发生紧急情况。

本节中列出的所有值都是固定值，Oracle ZFS Storage Appliance 作为一个设备，不提供调整这些参数的功能（也没有任何需要）。这些值作为实施详细信息，在此处提供仅供参考。随时可能更改，恕不另行通知。

注 - 为了避免在群集物理重新定位后造成数据损坏，应验证在新位置是否正确执行了所有群集布线。有关更多信息，请参见[“防止“记忆分裂”情况” \[163\]](#)

了解群集资源管理

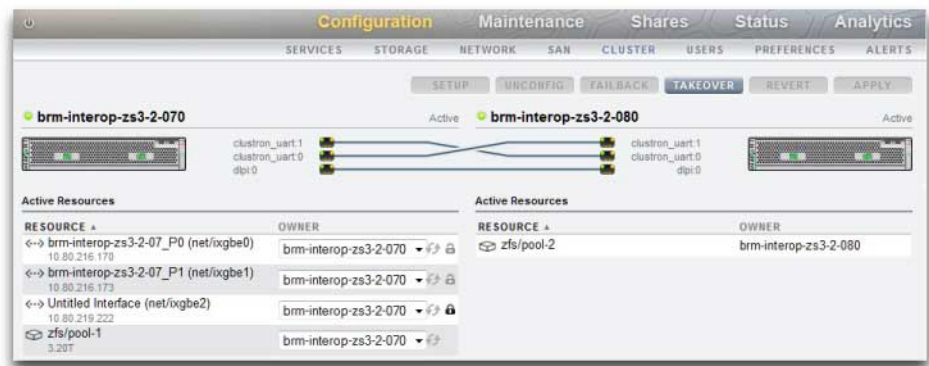
资源管理器负责确保激活正确的一组网络接口、适当的存储池处于活动状态，并且大量的配置参数在两个群集机头之间保持同步。该子系统的大多数活动都不会显示给管理员，但是会显示一个重要方面。资源分为几种类型，控制何时以及是否导入资源（使其处于活动状态）。请注意，活动状态的定义因资源类别而异。例如，网络接口属于网络类别，在激活接口后处于活动状态。三种最重要的资源类型是单件、专用和副本。

副本最简单：从不显示给管理员，也不显示在群集配置屏幕上（参见图 4）。副本始终存在，并且在两个机头上始终处于活动状态。通常，这些资源仅仅用作服务属性的容器，必须在两个机头之间保持同步。

与副本一样，单件资源也保持状态同步；但是，单件始终只在一个机头上处于活动状态。管理员可选择每个单件通常应该在哪个机头上处于活动状态；如果该机头发生故障，其对等设备将导入单件。单件是群集可用性特征的关键所在，一般认为单件资源从发生故障的机头移至未发生故障的对等设备，其中包括网络接口和存储池。由于网络接口是客户机用于查找一组已知存储服务的 IP 地址集合，因此将每个接口分配到的机头一定要与存储池客户机访问该接口的地址时看到的机头相同。在图 4 中，与 PrimaryA 接口关联的所有地址始终由导入 pool-0 的机头提供，而与 PrimaryB 关联的地址始终由与 pool-1 相同的机头提供。

专用资源仅为分配了该专用资源的机头所知晓，发生故障时不会被接管。这通常只对网络接口有用；请参见以下具体使用案例的论述。

图 10-4 ZS3-2 群集示例



还存在其他几种资源类型，这些是不会显示给管理员的实施详细信息。一种类型是共生体，指一个资源随另一个资源导入和导出。这种资源类型最重要的用途是表示存储池中的磁盘和闪存设备。这些资源称为磁盘集，必须始终在其包含的 ZFS 池之前导入。每个磁盘集包含外部存储附件中一半的磁盘；群集存储系统可连接任意数量的磁盘集（取决于硬件支持），每个 ZFS 池由一个或多个磁盘集中的存储设备构成。由于磁盘集可包含 ATA 设备，因此必须明确地将其导入和导出，以避免某些与从属关系相关的特定于多路径环境中所用 ATA 设备的行为。将磁盘表示成资源提供了一种在适当的时间执行这些活动的简便方法。管理员设置或更改存储池所有权时，将同时透明地更改与其关联的磁盘集的所有权分配。与所有共生体一样，磁盘集资源不会显示在群集配置用户界面中。

表 10-3 群集资源管理

资源	图标	普遍使用	发生故障时接管
SINGLETON		否	是
REPLICA	无	是	不适用
PRIVATE		否	否
SYMBIOTE	无	与父类型相同	与父类型相同

创建新资源后，最初将其分配给创建该资源的机头。除非该机头处于 AKCS_OWNER 状态，否则无法更改这种所有权。因此，有必要在通常应该拥有资源的机头上创建资源或在更改资源所有权之前接管。一般来说，从任一机头均可销毁资源，然而无法销毁导出的存储池。通常，无论分配的所有者是哪个机头，在当前控制资源的机头上将其销毁都能获得最佳效果。

大多数配置设置（包括服务属性、用户、角色、身份映射规则、SMB 自动主目录规则和 iSCSI 启动器定义）都将自动在两个机头上复制。因此，无论群集处于哪种状态，都不需要在两个机头上配置这些设置。如果一个设备在更改配置时发生故障，则在提供任何服务之前，下次引导设备而重新链接群集时，其配置将复制到另一个设备。存在少数几项例外：

- 无论通常将底层池分配给哪个机头，共享资源和 LUN 定义与选项都只能在控制该池的机头上进行设置。
- 不会复制“身份”服务的配置（即设备名称和位置）。
- 只能在分配了机箱的机头上看到为机箱指定的名称。
- 每个网络路由均绑定到特定接口。如果为每个机头分配的接口的地址在特定子网中，而且该子网包含设备应将流量定向到的路由器，则即便使用同一网关地址，也必须为每个此类接口创建一个路由。这样一来，控制两个机头之间的底层网络资源移动时，每个路由便可单独变为活动状态。有关更多详细信息，请参见“联网注意事项”。
- 不复制且从不共享 SSH 主机密钥。因此，如果未配置任何专用管理接口，则使用分配给发生故障的节点的地址登录 CLI 时会出现密钥不匹配的情况。用于访问 BUI 的 SSL 证书存在同样的限制。

因此，基本做法是透明地复制常见配置，并且管理员将为每个设备机头分配一组资源。这些资源分配进而将网络地址绑定到客户机应该看到的存储资源。无论哪个设备控制这组资源，客户机都能在预期的网络位置访问所需存储。

群集的接管和故障恢复

群集机头节点在任何指定时间均处于以下几种状态之一：

表 10-4 群集状态

状态	图标	CLI/BUI 表示	说明
UNCONFIGURED		Clustering is not configured (群集未配置)	没有任何群集的系统处于此状态。正在设置系统或从未完成群集设置任务。
OWNER		Active (takeover completed) (活动 (接管已完成))	已配置群集，该节点已控制群集中的所有共享资源。在用户界面中完成群集设置后，检测到对等设备已发生故障时（即接管后），系统将立即进入此状态。管理员手动执行故障恢复操作之前，它一直处于此状态。
STRIPPED		Ready (waiting for failback) (就绪 (等待故障恢复))	已配置群集，并且该节点不控制任何共享资源。在另一个节点的用户界面

状态	图标	CLI/BUI 表示	说明
			中完成群集设置后，或者在重新引导、电源断开或发生其他故障后，系统将立即进入 STRIPPED 状态。管理员手动执行故障恢复操作之前，节点一直处于此状态。
CLUSTERED		Active (活动)	已配置群集，并且两个节点均根据其资源分配拥有共享资源。如果每个节点都拥有 ZFS 池且处于 CLUSTERED 状态，则这两个节点构成的群集通常称为主动-主动群集。
-		Rejoining cluster ... (正在重新加入群集...)	设备近期已重新引导或设备管理软件在发生内部故障后正在重新启动。正在重新同步资源状态。
-		Unknown (disconnected or restarting) (未知 (已断开连接或正在重新启动))	对等设备已关闭电源或正在重新引导、其所有群集互连链路都已发生故障或尚未配置群集。

这些状态之间的转换发生在两项操作过程中：接管和故障恢复。

接管可在任何时间发生；如上所述，只要检测到对等设备故障就会尝试接管。也可使用群集配置 CLI 或 BUI 手动触发接管。这对于测试目的以及执行滚动软件升级（升级一个机头的同时，另一个机头提供运行较旧软件的服务，新软件经过验证后升级第二个机头）非常有用。最后，在引导一个机头且检测到其对等设备不存在时将进行接管。这样一来，一个机头发生永久性故障或两个机头暂时断电时，服务便可正常恢复。

故障恢复从不自动执行。修复并引导发生故障的机头后，它将重新链接群集（重新同步所有资源、资源属性和资源所有权的配置），并继续等待管理员执行故障恢复操作。在此之前，最初未发生故障的机头将继续提供所有服务。这样一来，便可在机头恢复生产服务之前全面调查最初触发接管的问题、验证新软件修订版或执行其他管理任务。由于故障恢复会对客户机造成中断，因此应根据特定于业务的需求和过程进行安排。存在一项例外：假设机头 A 发生了故障，机头 B 已接管。机头 A 重新链接群集后，如果检测到机头 B 不存在或已发生故障，机头 A 可以接管。原则是即使没有机会调查初始问题，提供服务总比不提供要好。因此，尽管故障恢复到之前发生故障的机头这项操作从不自动执行，但在任何时间依然都能执行接管。

设置群集时，初始状态包括启动设置的节点处于 OWNER（所有者）状态，另一个节点处于 STRIPPED 状态。执行初始故障恢复操作以将相应部分的共享资源交予 STRIPPED 节点后，两个节点都将处于 CLUSTERED 状态。如果两个群集节点都发生故障或关闭电源，则同时启动后将执行仲裁，其中一个将成为 OWNER（所有者）节点，另一个则成为 STRIPPED 节点。

故障恢复期间，所有外部资源（分配给对等设备的资源）都将导出，然后由对等设备导入。因发生故障而无法导入的池将触发 STRIPPED 节点的重新引导。如果尝试对发生故障的池进行故障恢复，会因导入失败而重新引导 STRIPPED 节点。

群集环境中的配置更改

绝大多数设备配置都以服务属性或共享资源/LUN 属性表示。共享资源和 LUN 属性随用户数据存储在存储池自身中（因而始终可由该存储资源的当前所有者访问），而服务配置存储在每个机头中。为了确保两个机头提供一致的服务，发生更改或之前发生故障的机头与其对等设备重新链接时，必须同步所有服务属性。由于所有服务都由副本资源表示，因此任何时间在任一机头上更改属性时，设备软件都会自动执行此同步。

因此，管理员没有必要（实在多余）复制配置更改。标准操作过程应该反映这一属性，要求在初始群集配置完成后只对两个机头中的一个进行更改。另请注意，在初始群集配置过程中，会将现有全部配置复制到新配置的对等设备上。因此，一般来说，我们得出群集配置更改的两个最佳做法：

- 在当前控制（或者将控制（如果将要创建新资源））底层存储或网络接口资源的机头上进行所有与存储和网络相关的配置更改。
- 在任一机头而非两个机头上进行其他所有更改。站点策略应指定将哪个机头视为该用途的主机头，进而应取决于正常运行的机头以及已配置的存储池数量。请注意，设备软件并不做此区分。

失忆是指进行不相交配置更改，随后在每个机头上丢失，而对等设备无法正常运行，这一问题在很大程度上有所夸大。对于 Oracle ZFS Storage Appliance 尤其如此，因为没有在每个机头上对系统配置进行独立更改的机制。这种简化大大降低了对集中配置系统信息库的需求，并且支持一种更为简单的方法：假设当前正在运行的机头具有正确的配置，其对等设备将在引导时与其同步。随着将来产品功能的增强，可以选择替代策略来解决配置差异，这种基本方法非常简单而且易于理解：第二个机头将采用现有生产系统已经使用的一组配置参数（因此极有可能正确）。为确保这种方法合适，管理员应确保发生故障的机头在修复后立即重新链接群集。

存储的群集注意事项

为了在群集中使用而对 Oracle ZFS Storage Appliance 进行大小调整时，另有两个非常重要的注意事项。或许最重要的决定是将所有存储池的所有权分配给同一机头还是在两个机头之间分割。此处有几点要加以权衡，如下表中所示。一般来说，应该在一个机头上配置池，但额定操作期间针对吞吐量优化或故障转移性能不容考虑时除外。处于故障转移状态时性能特征的确切变化在很大程度上取决于工作负荷的性质和大小。一般来说，某个机头沿任何特定轴提供的性能越接近最大值，工作负荷由该机头的对等设备接管时沿该轴的性能下降越明显。当然，如果包含多个池，两种工作负荷的性能都会下降。

请注意，在任一配置中，要使用任何 ReadZilla 设备，必须在分配了相应池所有权的机头上导入将设备分配到的池。也就是说，如果某个池因机头故障而被接管，即使导入该池的机头上也安装了未使用的 ReadZilla，也没有该池的读取缓存。出于这一原因，应[按照第 5 章 存储配置](#)文档中的说明配置主动-被动群集中的 ReadZilla。这不适用于 LogZilla 设备，LogZilla 位于存储结构中，始终可由已导入了池的机头访问。

表 10-5 存储的群集注意事项

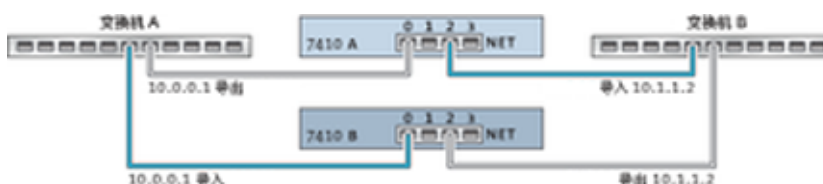
变量	单节点所有权	不同机头拥有多个池
总吞吐量（额定操作）	在任何时间最多可使用总 CPU 资源的 50%、DRAM 的 50% 和总网络连接的 50% 来提供服务。这种情况非常简单：只有一个机头处理客户机请求，因而另一个处于闲置状态。	在任何时间均可使用所有 CPU 和 DRAM 资源来提供服务。在任何时间最多可使用所有网络连接的 50%（需要在每个机头上安装备用网络设备来支持故障转移）。
总吞吐量（故障转移）	相对于额定操作，吞吐量无变化。	将使用未发生故障的机头的全部资源来提供服务。相对于额定操作，总吞吐量范围从大约 40% 到 100%，具体取决于额定操作期间的利用率。
I/O 延迟（故障转移）	ReadZilla 在故障转移操作期间不可用，对于适合可用的读取高速缓存的读取操作繁重的工作负荷，这会显著增加延迟。写入操作的延迟不受影响。	ReadZilla 在故障转移操作期间不可用，对于适合可用的读取高速缓存的读取操作繁重的工作负荷，这会显著增加延迟。读取和写入操作的延迟可能会因机头资源争用加剧而有所增加。原因是在未发生故障的机头上运行两种工作负荷而非通常的一种。每个机头上的额定工作负荷接近机头的最大能力时，故障转移状态下的延迟可能会非常高。
存储灵活性	共享资源和 LUN 可使用所有可用的物理存储。	特定池的共享资源和 LUN 只能使用分配给该池的存储。存储并不在池之间共享，因此如果一个池空间已满而另一个池具有可用空间，就会浪费某些存储。
网络连接	在每个机头提供服务的同时，可使用该机头上的所有网络设备。	在每个机头提供服务的同时，只能使用该机头上所有网络设备的一半。因此，每个池只能连接到物理上不相交的网络中的一半。

第二个重要的存储注意事项是采用无单点故障 (No Single Point of Failure, NSPF) 的池配置。由于使用群集意味着应用程序非常注重可用性，因此几乎没有充分的理由将存储池配置为允许单个 JBOD 发生故障导致可用性降低。这种方法的弊端是，NSPF 配置需要的 JBOD 数量多于单点故障配置，如果所需容量非常小，安装足够多的 JBOD 以在所需 RAID 级别实现 NSPF 可能并不经济。

网络的群集注意事项

网络设备、数据链路和接口发生故障时不会导致群集子系统机头也发生故障。为了保护系统免受网络故障（无论是在设备内部还是外部）的影响，应使用 IPMP 和/或 LACP。获取可用性的全面方法需要正确配置网络和全网范围内的冗余计划。

图 10-5 网络的群集



可以将网络接口配置为单件或专用资源，前提是网络接口具有静态 IP 配置。使用 DHCP 配置的接口必须是专用的，不建议在群集中使用 DHCP。如果将用于构建接口的所有数据链路和设备配置为单件资源，则它们一次都只能在一个机头上处于活动状态。同样，为了能够在故障转移状态下提供服务，也必须将每个机头上相应的设备连接到相同的网络。上图中显示了有关这一点的示例。

当您通过设备和数据链路构建网络接口时，要使群集能够正常操作，每个单件接口的两个机头上配备使用相同标识符和功能设备非常重要。由于设备标识符取决于设备类型以及 Sun 设备首次检测该设备时采用的顺序，因此群集机头必须安装完全相同的硬件。必须在两个机头上的每个插槽中插入完全相同的硬件，而且必须以相同的顺序在两个机头上的插槽中插入。合格的 Oracle 经销商或服务代表可协助您规划满足这些要求的硬件升级。

一个路由始终明确绑定到一个网络接口。路由在资源管理器中以共生体表示，且仅在其绑定到的接口运行时才能变为活动状态。因此，如果路由绑定到的接口当前处于备用模式（已导出），则在接管过程中激活该接口之前，该路由不起任何作用。配置了两个池且使其可用于通用子网时，这一点非常重要。如果子网是设备连接到一个或多个其他网络所使用的路由器所在的网络，则必须配置单独的路由（例如，第二个默认路由）并将其绑定到与该子网连接的每个活动接口和备用接口。

示例：

- 接口 e1000g3 分配给 "alice"，e1000g4 分配给 "bob"。
- 每个接口在 172.16.27.0/24 网络中都有一个地址，用于为可通过 172.16.27.1 连接的 172.16.64.0/22 网络中的客户机提供服务。
- 应该通过 172.16.27.1 创建两个到 172.16.64.0/22 的路由，一个应该绑定到 e1000g3，另一个绑定到 e1000g4。

较好的做法是，为每个群集机头分配一个只用于管理的 IP 地址（最有可能在专用管理网络上），并将此接口指定为专用资源。这样能够确保可从管理网络连接到正常运行的机头，即使该机头正处于 AKCS_STRIPPED 状态和等待故障恢复状态也是如此。这在正使用 LDAP 和 Active Directory 等服务并需要访问其他网络资源（当机头不提供服务）时非常重要。如果这样做并不切实可行，应将服务处理器连接到可靠的网络和/或串行终端集中器，以便能够使用系统控制台管理机头。

如果这两项操作均未执行，则直到故障恢复完成后才能管理或监视新引导的机头。您可能想要监视或管理当前为特定存储池提供服务的机头。当您要修改存储自身的某方面（例如修改共享资源属性或创建新 LUN）时，这一点很可能非常有用。要做到这一点，可使用服务接口之一执行管理任务或分配一个单独的单元接口，使其仅用于管理与其匹配的池。在这两种情况下，都应将其接口与其用于管理的池分配给同一机头。

专用本地 IP 接口

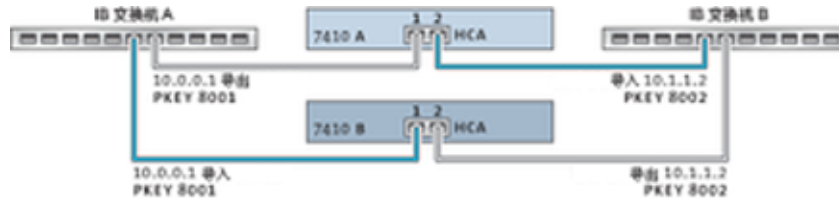
创建专用本地 IP 接口时，请遵循以下准则：

- 创建 IP 接口时，如果使用的名称与群集对等设备上的专用 IP 接口相同，则将在本地创建一个专用 IP 接口。
- 对等设备的专用接口使用的数据链路不可删除，且删除按钮变为灰色。
- 属于同一个 IPMP 组的 IP 接口都必须是相同的类型，且必须属于同一个机头。要创建一个 IPMP 组，您要么全部使用单元 IP 接口，要么全部使用专用 IP 接口，并且群集节点必须是这些接口的所有者。
- IPMP 组类型仅在创建组时设置，并取决于底层链路的类型。
- 属于 IPMP 组的 IP 接口不会显示在“群集：资源”页面上，因为要修改 IP 接口所有权，必须拥有 IPMP 组所有权。
- 专用 IPMP 组不会显示在“群集：资源”页面上，因为此所有权类型不可修改。

Infiniband 的群集注意事项

与在以太网设备基础上构建的网络一样，为确保免受设备内部和外部网络故障的影响，Infiniband 网络需要作为冗余结构拓扑的一部分。网络拓扑应包括 IPMP，以便通过更广泛的 HCA、交换机和子网管理器的冗余规划，保护系统免受链路级别网络故障的影响。

图 10-6 Infiniband 的群集注意事项

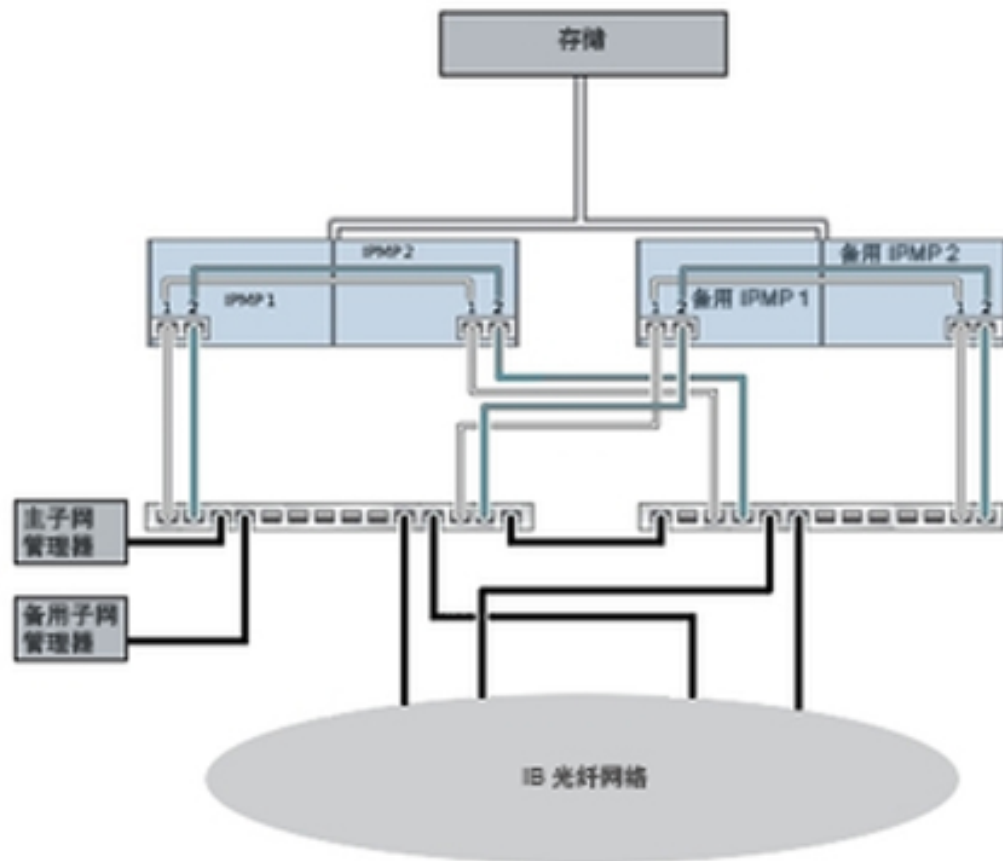


为确保正确的群集配置，必须在每个机头上完全相同的插槽中插入完全相同的 HCA。另外，必须在子网管理器上将每个相应的 HCA 端口配置到同一分区 (pkey) 中，使其具有完全相同的成员特权并连接到同一网络。为降低复杂性并确保适当的冗余，建议使每个端口仅属于 Infiniband 子网中的一个分区。可以将网络接口配置为单件或专用资源，前提是网络接口具有静态 IP 配置。如果将用于构建接口的所有 IB 分区数据链路和设备配置为单件资源，则它们在任何指定时间都只能在一个机头上处于活动状态。有关这一点，上图中显示了具体示例。对相应端口的分区成员身份的更改必须同时进行且遵循上述群集规则。合格的 Oracle 经销商或服务代表可协助您规划满足这些要求的硬件升级。

群集冗余路径方案

下图显示了子网管理器冗余的群集配置。将两个双端口 HCA 连接到一对冗余的服务器交换机可实现更大的冗余。

图 10-7 子网管理器冗余的群集配置



防止“记忆分裂”情况

群集系统中的一种常见故障模式称为记忆分裂；在这种情况下，每个群集机头都认为其对等设备已发生故障并尝试接管。这种情况不进行额外的逻辑推理，会导致大范围难以诊断或更正的破坏性意外行为。这种情况典型的起因是机头共享的通信介质发生故障；对于 Oracle ZFS Storage Appliance，群集 I/O 链路发生故障时会出现这种情况。除内置三重链路冗余（要避免触发接管，只需要一条链路）外，设备软件还会执行仲裁程序来确定哪个机头应该继续接管。

类似的产品采用多种仲裁机制，这些机制通常要求使用法定磁盘（使用 SCSI 预留空间）或法定服务器。为了支持使用 ATA 磁盘而无需其他硬件，Oracle ZFS Storage Appliance 采用一种不同的方法，依靠存储结构自身来提供所需的互斥性。仲裁过程包括尝试以预定义的顺序对存储结构中的每个可见 SAS 扩展器执行 SAS ZONE LOCK 命令。尝试获取所有这些锁定成功的设备将继续接管，另一个将对自身进行重置。由于引导并检测到对等设备无法连接的群集设备将尝试接管并开始相同的仲裁过程，因此将持续循环重置，直到至少恢复一条群集 I/O 链路为止。这可确保另一个机头以后的故障不会造成长时间中断。执行故障恢复后或处于 AKCS_OWNER 状态的机头最新续订其自己对存储结构的访问权限大约 10 秒后，即会解除这些 SAS 区域锁定。

这种仲裁机制非常简单、成本较低且无需其他硬件，但是要求两个群集设备都必须至少能够访问存储结构中的一个通用 SAS 扩展器。在正常情况下，每个设备均可访问所有扩展器，仲裁包括采用至少两个 SAS 区域锁定。但是，可能会出现多重故障情形，在这种情况下，设备无法访问任何通用扩展器。例如，如果拔下了两条 SAS 电缆或关闭了 JBOD 的电源，则每个设备可访问不相交的一部分扩展器。在这种情况下，每个设备都将成功锁定所有可连接的扩展器，断定对等设备已发生故障，并尝试继续接管。这样会因磁盘从属关系冲突和/或严重的数据损坏而导致不可修复的挂起。

请注意，尽管这种情况的后果非常严重，但是只会在多重故障情况（通常只在四重或更多故障的情况）下发生。嵌入 Oracle ZFS Storage Appliance 的群集解决方案目的在于确保不会发生单点故障，保护数据和可用性不受任何似乎合理的故障的影响，而不会造成系统的成本和复杂性过高。仍有可能发生大规模的多重故障而导致丢失服务和/或数据，就像没有 RAID 布局可保护系统免受无限数量磁盘故障的影响一样。

图 10-8 防止“记忆分裂”



幸运的是，大多数此类故障情形都由人为错误引起，可通过正确安装硬件和对员工进行群集设置和管理最佳做法方面的培训来全面预防。管理员始终都要确保三条群集 I/O 链路全部都已连接且正常运行（参见图），并按照设备随附的安装说明页中所示连接所有存储电缆。特别重要的一点是，将群集投入生产环境前以及此后必须能检测到每

个 JBOD 有两个路径（参见图），临时更改布线来支持容量增加或更换故障组件当然除外。管理员应使用警报来监视群集互连链路和 JBOD 路径的状态，并即时解决所有故障。通过确保维持正确的连接，可在硬件或软件组件发生故障时保证可用性和数据完整性。

图 10-9 群集的两个路径

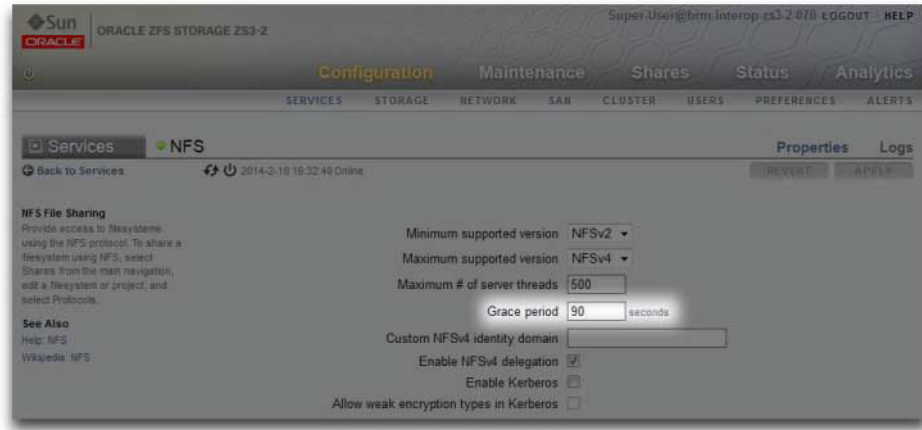


估计和降低接管影响

接管和故障恢复期间有一段时间间隔，在此期间客户机无法访问存储。这段时间间隔的长度因配置而异，并且对客户机的具体影响取决于客户机用于访问数据的协议。是否了解和降低这些影响关乎是成功地部署群集，还是在最糟糕的时刻发生故障而损失大量的资金。

NFS（所有版本）客户机通常不让应用程序软件发现中断，导致 I/O 操作延迟而服务器不可用。NFSv2 和 NFSv3 是无状态协议，在服务恢复后几乎能够立即恢复。NFSv4 在启动时有客户机宽限期，I/O 通常在此期间无法执行。此宽限期的持续时间可在 Oracle ZFS Storage Appliance 中进行调整（参见图），缩短这段时间可降低接管和/或故障恢复的明显影响。对于计划内中断，Oracle ZFS Storage Appliance 为 NFSv4 客户机提供无宽限期恢复，这样可避免宽限期延迟。有关无宽限期恢复的更多信息，请参见 NFS “属性” [185]。

图 10-10 群集宽限期



服务中断期间的 iSCSI 行为取决于启动器，但是如果服务在特定于客户机的超时期限内恢复，启动器通常会恢复。有关其他详细信息，请参见启动器的相关文档。iSCSI 目标通常能够在接管完成后立即提供服务，而不会出现进一步的延迟。

SMB、FTP 和 HTTP/WebDAV 是面向连接的协议。由于与这些服务关联的会话状态无法随底层存储和网络连接传输，因此使用其中一个协议的所有客户机都将在接管或故障恢复期间断开，并且必须在操作完成后重新连接。

尽管有多个因素会影响接管时间（以及与其密切相关的故障恢复时间），但是在大多数配置中，这些时间都由导入磁盘集资源所需的时间控制。每个磁盘集的导入时间范围通常为 15 到 20 秒，总时间与磁盘集数量线性相关。前面提到，磁盘集包括一个 JBOD 的一半，前提是这一半 JBOD 中的磁盘托架已填充且分配给存储池。未分配的磁盘和空磁盘托架对接管时间没有任何影响。导入磁盘集资源所需的时间不受可由管理员调整或更改的任何参数的影响，因此规划群集部署的管理员应执行以下任一操作：

- 限制安装的存储，以使客户机能够容许相关的接管时间，或者
- 调整超过预期最长接管时间的客户端超时值。

请注意，尽管磁盘集导入通常占接管时间的一大部分，但是这并不是唯一的因素。池导入过程中，必须重放所有意图日志记录，并且必须通过相应的服务共享每个共享资源和 LUN。对于单个共享资源或 LUN，执行这些活动所需的时间非常少（约几十毫秒），但是如果共享资源数量非常大，这会导致接管时间明显增加。因此，确保共享资源数量相对较少（几千甚至更少）可显著缩短这些时间。

对于任何指定配置，故障恢复时间通常都比接管时间长。这是因为故障恢复操作分为两个步骤：首先源设备导出其不是分配的所有者的所有资源，然后目标设备只对其自己分配的资源执行标准的接管过程。因此，如果发生了故障，从机头 A 故障恢复到机头 B

始终比从机头 B 接管机头 A 所需时间长。导出的磁盘集数量对较长的故障恢复时间的决定作用比对接管时间的决定作用小得多，因此确保共享资源和 LUN 数量较少对故障恢复的影响比对接管的影响要大。另请切记，故障恢复始终由管理员启动，而其导致的服务中断较长，因此可安排在适当的时间执行，确保由此造成的业务中断时间最短。

注意：本节引用的估计时间指的是软件/固件版本 2009.04.10,1-0。其他版本的执行方式可能有所不同，实际性能也可能不同。在生产环境中部署群集设备之前，测试接管及其对客户机应用程序的具体影响非常重要。

使用 BUI 进行群集配置

要配置或取消配置群集，请使用以下过程。

取消群集配置是一项破坏性操作，会使一个群集存储控制器恢复到出厂默认配置，并将所有资源的所有权重新分配给未发生故障的对等设备。取消群集配置的原因有以下两个：不再希望使用群集，而是希望配置两个单独的存储设备。要以新硬件或带有出厂全新设备软件的存储控制器更换发生故障的存储控制器（通常这种更换由服务提供商执行）。

▼ 配置群集

1. 将电源和至少一条以太网电缆连接到每个设备。
2. 按照下面“节点布线”部分的说明，将群集互连控制器通过电缆连接在一起。也可以继续群集设置，并在设置过程中动态地添加这些电缆。
3. 将 HBA 与共享 JBOD 通过电缆连接在一起，如设备随附的安装说明页中的 JBOD 布线图中所示。
4. 打开两个设备的电源，但是并不开始配置。从两个设备中任选一个要为其执行配置的设备。将其称为用于配置目的的主设备。连接到并访问该设备的串行控制台，在其上执行基于 tty 的初始配置，采用的方式与配置单机设备相同。注意：请勿在辅助设备执行基于 tty 的初始配置，群集设置期间将自动对其进行配置。
5. 在主设备上，进入 BUI 或 CLI 开始群集设置。如果已安装群集互连控制器，则可在初始设置过程中选择群集设置。或者，此时也可执行单机配置，稍后再执行群集设置。在后一种情况下，可以在 "Configuration" > "Cluster" 中单击 "Setup"（设置）按钮来执行群集配置任务。
6. 在群集设置的第一步，将显示活动群集链路图：您应该看到屏幕上显示三条蓝色实心电线，一条对应于一个连接。如果未看到，请立即添加缺少的电缆。看到全部三条电线后，即可单击 "Commit"（提交）按钮继续。

7. 输入第二个设备的设备名称和初始 root 用户密码（这相当于为新设备执行初始串行控制台设置）。单击 "Commit"（提交）按钮后，将在配置第二个设备时显示进度指示条。
8. 如果在主设备初始设置过程中设置群集，则此时系统将提示您执行初始配置，就像在单设备情况下一样。您所做的所有配置更改都将自动传播到另一个设备。继续初始配置，考虑以下限制和注意事项：通过 DHCP 配置的网络接口无法在机头之间进行故障转移，因此无法由客户机用来访问存储。所以，请确保为客户机要用来访问存储的所有网络接口分配静态 IP 地址。如果在基于 tty 的初始配置过程中选择了已配置 DHCP 的网络接口，并且希望将该接口用于客户机访问，则需要将其地址类型更改为静态，然后再继续。最佳做法包括为每个机头配置和分配用于管理的专用网络接口，这样，无论群集处于哪种状态，都将在网络（BUI 或 CLI）上通过任一机头启用管理。如果需要路由，请确保在要分配给每个机头的接口上创建路由。有关具体示例，请参见上一节。
9. 继续初始配置，直到配置存储池的步骤。执行接管时，群集对等设备可接管每个存储池以及客户机用于连接该存储池的网络接口。如果创建两个存储池，通常每个机头将为客户机提供对为其分配的池的访问权限；如果其中一个机头发生故障，另一个机头将为客户机提供对两个池的访问权限。如果创建一个池，仅当对等设备发生故障时，未分配池的机头才会为客户机提供服务。创建存储池时即会将其分配给机头；存储配置对话框中提供了用于单独创建分配给每个机头的池的选项。可以分配给池的最小存储单元是一个磁盘。如果创建了多个池，不要求它们的大小必须相同。注意，池越少，每个池的磁盘越多越好，因为这样可以简化管理并提高整体可用容量的百分比。强烈建议每个池都至少包括所有 JBOD 中的 8 个磁盘，越多越好。
10. 完成基本配置后，即可为每个机头分配资源。通常，只需分配网络接口，存储池已在存储配置步骤中自动分配。
11. 从群集用户界面中提交资源分配并执行初始故障恢复，具体说明如下。如果仍执行主设备的初始设置，将按设置顺序在最后显示该屏幕。如果要在初始设置后手动执行群集设置，请转到 "Configuration"（配置）/"Cluster"（群集）屏幕执行这些任务。有关详细信息，请参阅下面的群集用户界面。

▼ 取消群集配置

1. 选择要重置为出厂配置的存储控制器。请注意，如果更换发生故障的存储控制器，可跳至第 3 步，前提是发生故障的存储控制器无法在您的站点上恢复正常。
2. 从要重置为出厂配置的存储控制器的系统控制台上，恢复出厂设置。
3. 存储控制器将重置，其对等设备将正常开始接管。注意：在开始引导恢复出厂设置的存储控制器之前（即进展到超出引导菜单之前），应先关闭其电源，等待其对等设备完成接管。
4. 拔下群集互连电缆（参见上方），然后从群集的外部存储附件中拆离已关闭电源的存储控制器。

5. 在剩下的存储控制器上，在 "Configuration" (配置) > "Clustering" (群集) 屏幕上单击 "Unconfig" (取消配置) 按钮。所有资源都会分配给该存储控制器，并且该存储控制器不再是任何群集的成员。
6. 接下来可将拆离的存储控制器（如果有）连接到其自己的存储，然后打开电源并正常配置。如果更换发生故障的存储控制器，请将更换件连接到剩下的存储控制器和存储，然后开始上述群集设置任务。

注 - 如果群集包含两个或更多池，则取消配置后，所有池的所有权都将分配给剩下的存储控制器。在 2010.Q1.0.0 之前的软件版本中，不支持这种配置。如果运行较低的软件版本，必须执行以下操作之一：销毁一个或两个池，连接更换存储控制器，执行上述群集设置任务，并将一个池的所有权重新分配给更换存储控制器；或者升级到支持每个存储控制器包含多个池的 2010.Q1.0.0 或更高的软件版本。

使用 CLI 配置群集

▼ 关闭群集配置

1. 使用以下 CLI 命令验证群集状态：

```
nas-7420-1a:> configuration cluster
nas-7420-1a:configuration cluster> show
```

2. 下面是群集属性的示例："state" 表示运行命令的机头的状态，"peer_state" 表示另一个机头的状态。

```
state = AKCS_OWNER
description = Active (takeover completed)
peer_asn = 365ed33c-3b9d-c533-9349-8014e9da0408
peer_hostname = nas-7420-1b
peer_state = AKCS_STRIPPED
peer_description = Ready (waiting for failback)
```

3. 使用下表验证节点状态。

该节点	另一个节点	情况
AKCS_CLUSTERED	AKCS_CLUSTERED	两个节点都在正常情况下运行。

该节点	另一个节点	情况
AKCS_OWNER	AKCS_STRIPPED	该节点拥有所有资源，处于活动模式。另一个节点处于备用状态，没有资源。
AKCS_OWNER	重新引导	另一个节点正在重新引导，该节点拥有所有资源。
AKCS_OWNER	未知	该节点不知道伙伴节点。

注 - 如果机头的状态不一致，群集可能会遇到问题。请先联系 Oracle 支持人员，然后再继续。

▼ 关闭备用机头

1. 关闭备用机头，使用 CLI 运行以下命令：

```
nas-7420-1b:configuration cluster> cd /
nas-7420-1b:> maintenance system poweroff
This will turn off power to the appliance. Are you sure? (Y/N)
```

2. 要确认您要关闭另一个机头，请键入 "Y"。

注 - 如果两个机头都处于 AKCS_CLUSTERED 状态，将在未发生故障的机头上自动执行接管。

3. 确认备用机头已关闭，并且群集的状态为 "OWNER/unknown"。
4. 使用 CLI 运行以下命令来关闭活动机头：

```
nas-7420-1a:configuration cluster> cd /
nas-7420-1a:> maintenance system poweroff
This will turn off power to the appliance. Are you sure? (Y/N)
```

5. 要确认您要关闭活动机头，请键入 "Y"。
6. 确认两个机头已关闭电源。从 ILOM 提示符，运行以下命令：

```
-> show /SYS power_state
```

7. 关闭磁盘机框的电源。

▼ 取消群集配置

- 在 CLI 中取消配置群集等同于使用 BUI 取消配置按钮。如果用户尝试取消配置某个未处于正确状态的群集，则会显示错误消息。

```
configuration cluster> help
Subcommands that are valid in this context:

resources          => Configure resources

help [topic]       => Get context-sensitive help. If [topic] is specified,
                    it must be one of "builtins", "commands", "general",
                    "help", "script" or "properties".

show              => Show information pertinent to the current context

done              => Finish operating on "cluster"

get [prop]         => Get value for property [prop]. ("help properties"
                    for valid properties.) If [prop] is not specified,
                    returns values for all properties.

setup             => Run through initial cluster setup

failback          => Fail back all resources assigned to the cluster peer

takeover          => Take over all resources assigned to the cluster peer

unconfig          => Unconfigure the cluster

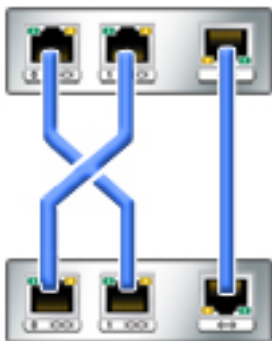
links             => Report the state of the cluster links
```

群集节点布线

必须使用位于控制器背面的群集互连端口将群集机头节点连接在一起。

ZS3-2 群集布线

图 10-11 ZS3-2 群集布线



ZS3-2 控制器提供了三条冗余链路来支持机头通信：两个串行链路（前两个连接器）和一条以太网链路（第三个连接器）。

使用直通 Cat 5 或更好的以太网电缆（群集配置随附三条 1 米的电缆），按照左侧的图连接机头节点。

群集布线可在打开任一机头节点电源之前执行，也可在执行群集设置指导性任务时实时执行。用户界面中将显示每条链路的状态，如本页中后面所示。必须先建立全部三条链路，然后才能继续群集配置。

ZS3-4 和 7x20 群集布线

图 10-12 ZS3-4 和 7x20 群集布线



ZS3-4 和 7x20 控制器提供了三条冗余链路来支持机头通信：两个串行链路（外面的两个连接器）和一条以太网链路（中间的连接器的）。

使用直通 Cat 5 或更好的以太网电缆（群集配置随附三条 1 米的电缆），按照左侧的图连接机头节点。

群集布线可在打开任一机头节点电源之前执行，也可在执行群集设置指导性任务时实时执行。用户界面中将显示每条链路的状态，如本页中后面所示。必须先建立全部三条链路，然后才能继续群集配置。

存储机框布线

开始群集配置之前，需要将存储机框连接到两个设备。请参见《[Oracle ZFS Storage Appliance 安装指南](#)》中的“安装”或者按照随系统提供的快速设置海报操作。

群集配置 BUI 页面

"Configuration"（配置）> "Cluster"（群集）视图提供了群集卡状态、群集机头节点状态 and 所有资源的图形概览。

图 10-13 配置群集视图



此界面包含以下对象：

- 每个系统的缩略图，正在访问其管理接口的系统显示在左侧。每个缩略图都标有规范的设备名称及其当前的群集状态（上面的图标和描述性标签）。
- 随硬件动态更新的每个群集卡连接的缩略图：某条链路已连接且处于活动状态时，将以一条实线连接该链路；该连接已断开或其他系统正在重新启动/重新引导时，这条线将消失。
- 当前分配给每个系统的专用和单件资源的列表（参见上方的“简介”），这些资源连同资源的各种属性以列表的形式显示在每个群集节点的缩略图下方。
- 对于每个资源，将该资源分配到的设备（即两个设备都处于 CLUSTERED 状态时提供资源的设备）。当前设备处于 OWNER（所有者）状态时，owner（所有者）字段将作为弹出式菜单显示，您可对其进行编辑，然后单击 "Apply"（应用）提交。

- 对于每个资源，表示其是否为专用资源的锁定图标。当前设备处于 OWNER（所有者）或 CLUSTERED 状态时，依次单击锁定图标和 "Apply"（应用）可将资源锁定到该设备（使其成为专用资源）或者解除锁定（使其成为单件资源）。请注意，属于远程对等设备的专用资源不会显示在任一资源列表中。

BUI 中包含以下按钮：

表 10-6 机框布线界面按钮

按钮	说明
Setup（设置）	如果尚未配置群集，可执行群集设置指导性任务，然后返回当前屏幕。有关该任务的详细说明，请参见上文。
Unconfig（取消配置）	通过取消群集配置，将节点升级到单机操作。有关该任务的详细说明，请参见下文。
Apply（应用）	如果资源修改待处理（以黄色突出显示的行），则将这些更改提交至群集。
Revert（还原）	如果资源修改待处理（以黄色突出显示的行），则恢复这些更改并显示当前的群集配置。
Failback（故障恢复）	如果当前设备（左侧）处于 OWNER（所有者）状态，则将另一个设备拥有的资源故障恢复到该设备，使两个节点都保持 CLUSTERED 状态（主动/主动）。
Takeover（接管）	如果当前设备（左侧）处于 CLUSTERED 或 STRIPPED 状态，则强制另一个设备重新引导并接管其资源，使当前设备变为 OWNER（所有者）状态。

◆◆◆ 第 11 章

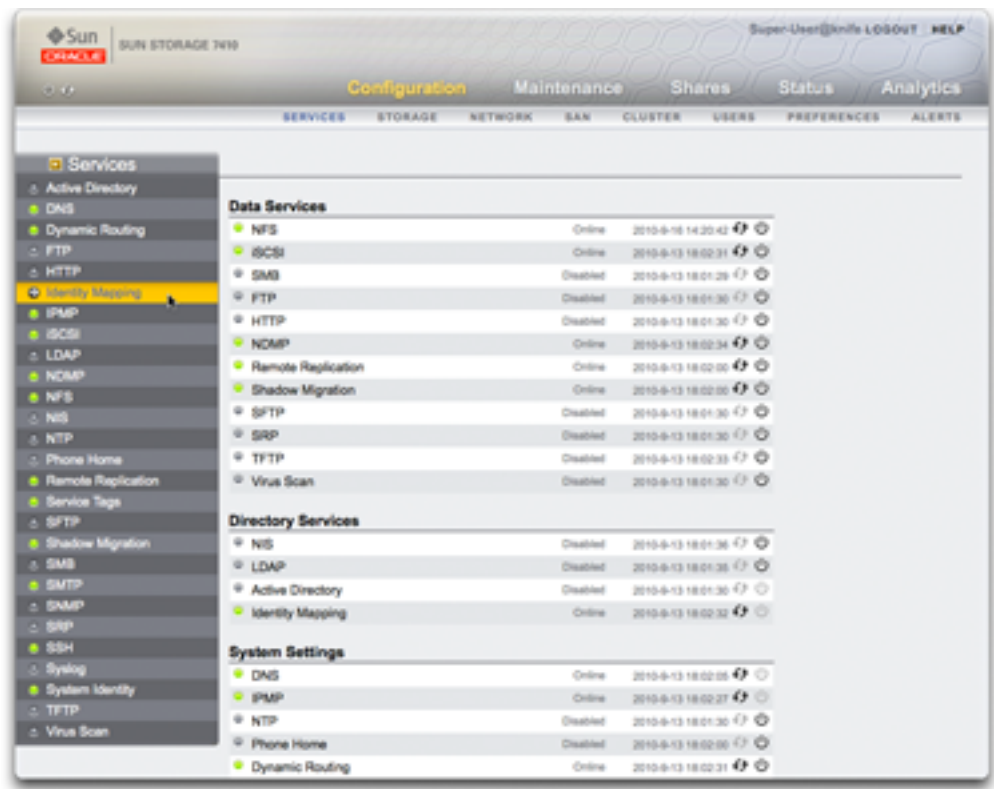
ZFSSA 服务

本服务屏幕的功能是作为各项服务之间的快速导航侧面板。

可用服务

您可以配置以下 ZFSSA 服务：

图 11-1 服务配置 BUI 页面



数据服务

表 11-1 可用数据服务

服务	说明	使用的端口
“NFS” [185]	通过 NFSv3 和 NFSv4 协议的文件系统访问	111 和 2049
“iSCSI” [189]	通过 iSCSI 协议的 LUN 访问	3260 和 3205
“SMB” [191]	通过 SMB 协议的文件系统访问	SMB-over-NetBIOS 139
SMB-over-TCP 445		
NetBIOS 数据报 138		

服务	说明	使用的端口
NetBIOS 名称服务 137		
“FTP” [206]	通过 FTP 协议的文件系统访问	21
“HTTP” [208]	通过 HTTP 协议的文件系统访问	80
“NDMP” [209]	NDMP 主机服务	10000
“远程复制” [216]	远程复制	216
“影子迁移” [216]	影子数据迁移	
“SFTP” [216]	通过 SFTP 协议的文件系统访问	218
“SRP” [219]	通过 SRP 协议的块访问	
“TFTP” [219]	通过 TFTP 协议的文件系统访问	
“病毒扫描” [220]	文件系统病毒扫描	

目录服务

注意：0-99 的 UID 和 GID 由操作系统供应商预留用于将来的应用程序。分层产品的最终系统用户或供应商使用它们不受支持并且可能会导致将来的应用程序出现安全相关问题。

表 11-2 可用目录服务

服务	说明	使用的端口
“NIS” [223]	通过 NIS 服务验证用户和组	
“LDAP” [225]	通过 LDAP 目录验证用户和组	389
“Active Directory” [228]	通过 Microsoft Active Directory Server 验证用户	
“身份映射” [233]	在 Windows 实体与 Unix ID 之间进行映射	

服务设置

表 11-3 服务设置

服务	说明	使用的端口
“DNS” [239]	域名服务客户机	53
“动态路由” [241]	RIP 和 RIPng 动态路由协议	

服务	说明	使用的端口
“IPMP” [242]	用于 IP 故障转移的 IP 多路径	
“NTP” [243]	网络时间协议客户机	
“回拨” [246]	产品注册和支持配置	443
“服务标签” [249]	产品清单支持	443
“SMTP” [250]	配置外发邮件服务器	
“SNMP” [251]	通过 SNMP 发送关于警报的陷阱并提供设备状态信息	
“Syslog” [254]	通过 syslog 转发发送关于警报的 syslog 消息并转发服务 syslog 消息	
“系统标识” [258]	系统名称和位置	

远程访问服务

表 11-4 可用远程访问服务

服务	说明	使用的端口
“SSH” [259]	通过 SSH 访问 CLI	22
“REST” [249]	RESTful API	

安全服务

表 11-5 可用安全服务

服务	说明	使用的端口
Kerberos	Kerberos V 验证	88
Kerberos V 更改和设置密码 (SET_CHANGE)	464	
Kerberos V 更改和设置密码 (RPCSEC_GSS)	749	

所需的最少端口数

要提供网络安全，可以在网络体系结构中部署防火墙。端口号用于创建防火墙规则并通过指定主机和服务在网络上唯一标识事务。

以下列表显示了创建防火墙规则（以使设备发挥全部功能）所需的最少端口数：

传入端口

- icmp/0-65535 (PING)
- tcp/1920 (EM)
- tcp/215 (BUI)
- tcp/22 (SSH)
- udp/161 (SNMP)

传出端口

- tcp/80 (WEB)
- tcp/443 (SSL WEB)

注 - 传出端口 tcp/443 用于发送回拨消息、上载支持包以及更新通知。对于复制，尽可能使用通用路由封装 (Generic Routing Encapsulation, GRE) 隧道。这样可以让流量在后端接口上运行，并避免在可能减慢流量的地方设置防火墙。如果 GRE 隧道在 NFS 核心上不可用，则您必须在前端接口运行复制。在这种情况下，端口 216 还必须是开放的。

使用 BUI 配置服务

您可以使用 BUI 服务屏幕查看和修改以上各表中所述的服务和设置。双击某个服务行可查看该服务的定义屏幕。下表介绍了服务屏幕中的图标和按钮：

表 11-6 服务 BUI 页面图标和按钮

图标	说明
	转到服务屏幕以配置属性和查看日志。将鼠标悬停在某个服务上时会出现此按钮
	服务已启用且正常工作。
	服务已脱机或已禁用。
	服务有问题，操作员需要注意。
	启用或禁用服务
	重新启动服务

图标	说明
	无法启用/禁用此服务
	重新启动当前不可用的服务。您必须首先启用该服务。

▼ 查看特定服务屏幕

1. 要查看或编辑某个特定服务的属性，请将鼠标悬停在位于服务名称左侧的服务状态图标上。
2. 状态图标将变为一个箭头图标，您可以单击该图标来显示所选服务的属性屏幕。

▼ 查看特定服务屏幕

- 在任何服务屏幕中，您都可以通过单击 "Services"（服务）标题（位于每个屏幕的左上角附近）左侧的小箭头图标来显示包含所有服务的侧面板。再次单击该图标可隐藏列表。

▼ 启用服务

- 如果某项服务未联机，可单击电源图标  使该服务联机 

▼ 禁用服务

- 如果某项服务已联机并希望将其禁用，可单击电源图标  使该服务脱机 

▼ 定义属性

1. 要定义服务的属性，请双击服务。
2. 更改属性并单击 "APPLY"（应用）按钮。

- 3. 要重置属性，请单击 "REVERT"（恢复）。

▼ 查看服务日志

- 1. 某些服务还提供了服务日志，日志中包含有助于诊断服务问题的信息。如果服务屏幕的右上角有 "Logs"（日志）按钮，说明该服务提供有日志。日志可提供以下信息：
 - 服务状态的更改时间
 - 来自服务的错误消息
- 2. 日志内容特定于各项单独的服务，并且会随设备软件将来的更新而更改。下面是此版本的设备中常用消息的示例：

日志消息示例	说明
Executing start method（正在执行启动方法）	服务正在启动
Method "start" exited with status 0（方法 "start"（启动）退出并显示状态 0）	服务报告了一次成功的启动（0 == 成功）
Method "refresh" exited with status 0（方法 "refresh"（刷新）退出并显示状态 0）	服务根据其服务设置成功刷新了其配置
Executing stop method（正在执行停止方法）	服务正在关闭
Enabled（已启用）	已检查服务状态来查看是否应将其启动（例如在系统引导期间），并且发现其处于已启用状态
Disabled（已禁用）	已检查服务状态来查看是否应将其启动（例如在系统引导期间），并且发现其处于已禁用状态

使用 CLI 配置服务

CLI 服务部分位于 configuration services 下。使用 show 命令列出所有服务的当前状态：

以下示例来自“NTP” [243]服务：

```
[ Oct 11 21:05:31 Enabled. ]
[ Oct 11 21:07:37 Executing start method (...). ]
[ Oct 11 21:13:38 Method "start" exited with status 0. ]
```

示例中的第一个日志事件表明系统在 21:05 引导。第二个条目记录了服务在 21:07:37 开始启动，且在 21:13:38 完成启动。由于 NTP 的性质和系统时钟调整，此服务可能要花费几分钟的时间才能完成启动，如日志所示。

```
caji:> configuration services
caji:configuration services> show
Services:
    ad => disabled
    smb => disabled
    dns => online
    dynrouting => online
    ftp => disabled
    http => disabled
    identity => online
    idmap => online
    ipmp => online
    iscsi => online
    ldap => disabled
    ndmp => online
    nfs => online
    nis => disabled
    ntp => disabled
    replication => online
    scrk => disabled
    sftp => disabled
    shadow => online
    smtp => online
    snmp => disabled
    ssh => online
    syslog => disabled
    tags => online
    tftp => disabled
    vscan => disabled

Children:
    ad => Configure Active Directory
    smb => Configure SMB
    dns => Configure DNS
    dynrouting => Configure Dynamic Routing
    ftp => Configure FTP
    http => Configure HTTP
    identity => Configure System Identity
    idmap => Configure Identity Mapping
    ipmp => Configure IPMP
    iscsi => Configure iSCSI
    ldap => Configure LDAP
    ndmp => Configure NDMP
    nfs => Configure NFS
    nis => Configure NIS
    ntp => Configure NTP
    replication => Configure Remote Replication
    scrk => Configure Phone Home
    sftp => Configure SFTP
    shadow => Configure Shadow Migration
    smtp => Configure SMTP
    snmp => Configure SNMP
    srp => Configure SRP
    ssh => Configure SSH
    syslog => Configure Syslog
    tags => Configure Service Tags
```

```
tftp => Configure TFTP
vscan => Configure Virus Scan
routing => Configure Routing Table
```

▼ 选择服务

1. 选择某项服务后，您可以查看其状态、将其启用或禁用以及设置其属性。
2. 通过输入服务名称来选择服务。例如，要选择 **nis**：

```
caji:configuration services> nis
caji:configuration services nis>
```

▼ 查看服务状态

- 可以使用 **show** 命令查看服务状态：

```
caji:configuration services nis> show
Properties:
    <status> = online
    domain = fishworks
    broadcast = true
    ypservers =
```

▼ 启用服务

- 可以使用 **enable** 命令启用服务：

```
caji:configuration services nis> enable
```

▼ 禁用服务

- 可以使用 **disable** 命令禁用服务：

```
caji:configuration services nis> disable
```

▼ 设置属性

1. 可以使用 **set** 命令设置所选服务的属性。
2. 在设置属性后，可以使用 **commit** 命令保存并激活新配置：

```
caji:configuration services nis> set domain="mydomain"
      domain = mydomain (uncommitted)
caji:configuration services nis> commit
caji:configuration services nis> show
Properties:
      <status> = online
      domain = mydomain
      broadcast = true
      ypservers =
```

3. 注意：属性名称类似于它们在 BUI 中的名称，但 CLI 名称通常较短且有时是缩写。

▼ 查看服务帮助

- 可以键入 **help** 来查看适用于服务的所有命令：

```
caji:configuration services nis> help
Subcommands that are valid in this context:

help [topic]          => Get context-sensitive help. If [topic] is specified,
                        it must be one of "builtins", "commands", "general",
                        "help", "script" or "properties".

show                  => Show information pertinent to the current context

commit                => Commit current state, including any changes

done                  => Finish operating on "nis"

enable                => Enable the nis service

disable               => Disable the nis service

get [prop]            => Get value for property [prop]. ("help properties"
                        for valid properties.) If [prop] is not specified,
                        returns values for all properties.

set [prop]            => Set property [prop] to [value]. ("help properties"
                        for valid properties.) For properties taking list
                        values, [value] should be a comma-separated list of
                        values.
```


NFS

网络文件系统 (Network File System, NFS) 是一种行业标准协议，用于通过网络共享文件。Sun ZFS Storage Appliance 支持 NFS 版本 2、3 和 4。有关如何构建文件系统名称空间的详细信息，请参见“文件系统名称空间” [271] 部分。有关本地用户的 NFS 的信息，请参见第 7 章 用户配置。

属性

- Minimum supported version (支持的最低版本) – 使用此下拉列表来控制设备支持哪些 NFS 版本。
- Maximum supported version (支持的最高版本) – 使用此下拉列表来控制设备支持哪些 NFS 版本。
- Maximum # of server threads (最大服务器线程数) – 定义并发 NFS 请求的最大数目 (从 20 到 1000)。这至少应当涵盖您预期的并发 NFS 客户机的数目。
- Grace period (宽限期) – 定义从计划外中断重新引导设备 (从 15 秒到 600 秒) 后所有客户机必须在多少秒内恢复锁定状态。该属性只影响 NFS v4 客户机 (NFS v3 是无状态协议，因此没有要恢复的状态)。在此期间内，NFS 服务只处理旧锁定状态的回收。在宽限期结束之前，不会处理对服务的其他请求。默认宽限期为 90 秒。减小宽限期将使得 NFS 客户机在服务器重新引导之后能够更快地恢复操作，但也会增加客户机无法恢复其所有锁定状态的可能性。在计划内中断期间，Oracle ZFS Storage Appliance 为 NFSv4 客户机提供无宽限期锁定状态恢复。在以下事件过程中出现计划内中断，例如《Oracle ZFS Storage Appliance 客户服务手册》中的“更新”、使用 CLI maintenance system reboot 命令重新引导设备或使用 BUI 电源图标  重新引导。对于计划内中断，NFS 服务会处理对服务的所有请求，而不会造成宽限期延迟。
- Custom NFSv4 identity domain (定制 NFSv4 标识域) – 使用此属性定义用于映射 NFSv4 用户和组标识的域。如果您没有设置此属性，则设备将使用 DNS 来获取标识域：首先检查 _nfsv4idmapdomain DNS 资源记录，然后回退到 DNS 域自身。
- Enable NFSv4 delegation (启用 NFSv4 委托) – 选择此属性将允许客户机在本地缓存文件且无需联系服务器即可进行修改。此选项默认情况下启用且通常可提高性能，但在极少数的情况下可能会引起问题。如果要禁用此设置，只能在对特定工作负荷进行仔细的性能测量并验证了这样设置具有相当的性能优势后进行。此选项只影响 NFSv4 挂载。
- Mount visibility (挂载可见性) – 此属性允许您限制有关 NFS 客户机共享资源访问列表和远程挂载等信息的可用性。Full allows full access (完全允许完全访问)。受限的限制访问，比如客户机仅能查看允许其访问的共享资源。客户机看不到在服务器上定义的共享资源或服务器中其他客户机完成的远程挂载的访问列表。默认情况下，此属性设置为 "Full" (完全)。
- Enable Kerberos (启用 Kerberos) – 启用/禁用 Kerberos 服务。

- * Allow weak encryption types in Kerberos (Kerberos 中允许的脆弱加密类型)
 - 启用/禁用对 DES (des-cbc-crc, des-cbc-md5) 以及可使用 HMAC/md5 导出的 ArcFour(arcfour-hmac-exp) 的支持。默认情况下, 该属性处于禁用状态。
- * Kerberos realm (Kerberos 领域) – 领域是一个逻辑网络, 类似于域, 它定义了位于同一个主 KDC 下的一组系统。领域名称可以包括任意 ASCII 字符串。通常, 领域名与 DNS 域名相同, 不同的是领域名采用大写。使用此约定可帮助您区分 Kerberos 服务发生的问题与 DNS 名称空间发生的问题, 同时仍使用熟悉的名称。
- * Kerberos master KDC (Kerberos 主 KDC) – 在每个领域中, 必须有一台维护主体数据库的主副本的服务器。主 KDC 与从 KDC 之间的最大区别是只有主 KDC 能够处理数据库管理请求。例如, 必须在主 KDC 上更改密码或添加新主体。
- * Kerberos slave KDC (Kerberos 从 KDC) – 从 KDC 包含主体数据库的重复副本。主 KDC 服务器和从 KDC 服务器都可创建用于建立验证的票证。
- * Kerberos admin principal (Kerberos 管理主体) – 此属性用于标识管理员。按照约定, 主体名称分为三个部分: 主元素、实例和领域。您可以将主体指定为 joe、joe/admin 或 joe/admin@ENG.EXAMPLE.COM。此属性仅用于设置系统的 Kerberos 服务主体, 不可保留。
- * Kerberos admin password (Kerberos 管理密码) – 定义管理员密码。此属性仅用于设置系统的 Kerberos 服务主体, 不可保留。
- Oracle Intelligent Storage Protocol (Oracle 智能存储协议) – NFSv4 服务支持 Oracle Intelligent Storage Protocol (Oracle 智能存储协议), Oracle Database NFSv4 客户机可通过该协议将优化信息传递到 ZFS Storage Appliance NFSv4 服务器。有关更多信息, 请参阅“[Oracle Intelligent Storage Protocol](#)” [428]。

更改服务属性记录在“[使用 BUI 配置服务](#)” [179]和“[使用 CLI 配置服务](#)” [181]中。

将 NFS 最低版本和最高版本设置为相同的值将导致设备仅与使用该版本的客户机进行通信。如果您发现与一个 NFS 版本或另一个版本有关的问题 (例如, 在您的工作负荷下某个 NFS 版本的性能特征), 并且您希望强制客户机仅使用工作性能最好的版本, 则这非常有用。

Kerberos Realms (Kerberos 领域)

配置 Kerberos 领域将创建特定的服务主体并向系统的本地密钥表中添加必要的密钥。必须在配置 Kerberized NFS 前配置“[NTP 服务](#)” [243]。为支持基于 Kerberos 的 NFS, 将创建并更新以下服务主体:

```
host/node1.example.com@EXAMPLE.COM
nfs/node1.example.com@EXAMPLE.COM
```

如果将设备组成群集, 会生成各个群集节点的主体和密钥:

```
host/node1.example.com@EXAMPLE.COM
nfs/node1.example.com@EXAMPLE.COM
host/node2.example.com@EXAMPLE.COM
```

nfs/node2.example.com@EXAMPLE.COM

如果已经创建了这些主体，则配置领域将重置这些主体中每个主体的密码。如果您将设备配置为加入 Active Directory 域，则无法将其配置为 Kerberos 领域的一部分。

有关设置 KDC 和基于 Kerberos 的客户机的信息，请参见 http://docs.oracle.com/cd/E26502_01/html/E29015/index.html (http://docs.oracle.com/cd/E26502_01/html/E29015/index.html)。在设置 Kerberos 的 NFS 属性后，在 "Shares" (共享) > "Filesystem" (文件系统) > "Protocols" (协议) 屏幕上将安全模式设置为使用 Kerberos 的模式。

设备将以下端口用于 Kerberos。

- Kerberos V 验证：88
- Kerberos V 更改和设置密码 SET_CHANGE：464
- Kerberos V 更改和设置密码 RPCSEC_GSS：749

注意：基于 Kerberos 的 NFS 客户机必须使用解析为相应主体的 FQDN 的 IP 地址访问设备。例如，如果设备配置有多个 IP 地址，只有解析为设备的 FQDN 的 IP 地址可由基于 Kerberos 的 NFS 客户机使用。

服务日志

与 NFS 服务相关的日志包括：

表 11-7 可用于 NFS 的日志

日志	说明
network-nfs-server:default	主 NFS 服务器日志
appliance-kit-nfsconf:default	设备 NFS 配置事件的日志
network-nfs-cbd:default	NFSv4 回调守护进程的日志
network-nfs-mapid:default	NFSv4 mapid 守护进程的日志 – 该守护进程映射 NFSv4 用户和组凭证
network-nfs-status:default	NFS statd 守护进程的日志 – 该守护进程有助于执行 NFS 锁定的崩溃和恢复功能
network-nfs-nlockmgr:default	NFS lockd 守护进程的日志 – 该守护进程支持文件的记录锁定操作

NFS Analytics

在《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“Analytics”部分中可以监视 NFS 活动。这包括：

- 每秒 NFS 操作数
- ... 按操作类型 (读取/写入/...)
- ... 按共享资源名称
- ... 按客户机主机名
- ... 按访问的文件名
- ... 按访问延迟

注意：当 NFS 服务器重新引导或故障转移时，文件名在服务器上是不知的，直至从客户机上重新打开。文件在 Analytics 工作表中显示为未知。

NFS BUI 和 CLI 属性

下表介绍了 CLI 属性与上述 BUI 属性之间的映射。

表 11-8 NFS BUI 和 CLI 属性

CLI 属性	BUI 属性
version_min	Minimum supported version (支持的最低版本)
version_max	Maximum supported version (支持的最高版本)
nfsd_servers	Maximum # of server threads (最大服务器线程数)
grace_period	Grace period (宽限期)
mapid_domain	Custom NFSv4 identity domain (定制 NFSv4 标识域)
enable_delegation	Enable NFSv4 delegation (启用 NFSv4 委托)
mount_visibility	Client share information restriction level (客户机共享资源信息限制级别)
krb5_allow_weak_crypto	Permits weak encryption types (arcfour-hmac-md5-exp, des-cbc-md5, and des-cbc-crc) in Kerberos (Kerberos 中允许的脆弱加密类型 (arcfour-hmac-md5-exp、des-cbc-md5 和 des-cbc-crc))
krb5_realm	Kerberos Realm (Kerberos 领域)
krb5_kdc	Kerberos master KDC (Kerberos 主 KDC)
krb5_kdc2	Kerberos slave KDC (Kerberos 从 KDC)
krb5_admin	Kerberos admin principal (Kerberos 管理主体)

▼ 通过 NFS 共享文件系统

1. 转到 "Configuration" (配置) -> "Services" (服务) 屏幕。
2. 检查 NFS 服务是否已启用且处于联机状态。如果不是，请启用该服务。
3. 转到[第 12 章 共享资源、项目和模式](#)屏幕，然后编辑现有共享资源或创建新共享资源。
4. 单击您编辑的共享资源的 "Protocols" (协议) 选项卡并检查是否已启用 NFS 共享。您还可以在此屏幕上配置 NFS 共享模式 (读取/读取+写入)。

iSCSI 服务

在设备上配置 LUN 时，您可以通过 Internet 小型计算机系统接口 (Internet Small Computer System Interface, iSCSI) 目标导出该卷。通过 iSCSI 服务，iSCSI 启动器可以使用 iSCSI 协议访问目标。

该服务支持使用 iSNS 协议进行搜索、管理和配置。iSCSI 服务支持使用 CHAP 的单向验证 (目标对启动器进行验证) 和双向验证 (目标和启动器相互验证)。此外，该服务还支持在 RADIUS 数据库中进行 CHAP 验证数据管理。

系统首先执行验证，然后进行授权，这在两个独立的步骤中进行。

注 - 有关配置 iSCSI 启动器和目标的示例，请参见[第 6 章 存储区域网络配置](#)部分。

iSCSI 服务属性

表 11-9 iSCSI 服务属性

属性	说明
Use iSNS (使用 iSNS)	是否启用 iSNS 搜索
iSNS Server (iSNS 服务器)	iSNS 服务器
Use RADIUS (使用 RADIUS)	是否启用 RADIUS
RADIUS Server (RADIUS 服务器)	RADIUS 服务器
RADIUS Server Secret (RADIUS 服务器密钥)	RADIUS 服务器的密钥

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

iSCSI 验证服务

如果本地启动器具有 CHAP 名称和 CHAP 密钥，则系统将执行验证。如果本地启动器没有 CHAP 属性，则系统不会执行任何验证，因此所有启动器都将符合授权条件。

iSCSI 服务授权

通过 iSCSI 服务，您可以指定能够在启动器组内使用的启动器的全局列表。

iSCSI 服务目标和启动器

有关 iSCSI 目标和启动器的详细信息，请参见[第 6 章 存储区域网络配置](#)。

iSCSI 故障排除

如果您的启动器无法连接到目标，请执行以下操作：

- 确保启动器的 IQN 与在启动器列表中标识的 IQN 匹配。
- 检查 iSNS 服务器的 IP 地址是否正确以及是否配置了 iSNS 服务器。
- 检查目标的 IP 地址在启动器端是否正确。
- 检查启动器 CHAP 名称和密钥在两端是否相同。
- 确保目标 CHAP 名称和密钥与任何启动器的名称和密钥都不相同。
- 检查 RADIUS 服务器的 IP 地址和密钥是否正确，以及是否配置了 RADIUS 服务器。
- 检查访问 LUN 的启动器是否为该 LUN 的启动器组的成员。
- 检查导出该 LUN 的目标是否处于联机状态。
- 检查 LUN 的运行状态是否为联机。
- 检查每个 LUN 的逻辑单元号。

在故障转移/故障恢复期间，如果来自 Red Hat 客户机的 iSER Reduced Copy I/O 未能避免发生故障，请执行以下操作：

- 在 `/etc/iscsi/iscsid.conf` 文件中，将 `node.session.timeo.replacement_timeout` 参数修改为 300sec。

SMB 服务

通过 SMB 服务可以使用 SMB 协议访问文件系统。支持的 SMB 版本为：SMB1、SMB2.0。必须从第 12 章 共享资源、项目和模式配置中将文件系统配置为使用 SMB 进行共享。

SMB 服务属性

- LAN Manager compatibility level (LAN Manager 兼容性级别) – 支持的验证模式 (LM、NTLM、LMv2、NTLMv2)。有关各兼容性级别中支持的验证模式的详细信息，请参阅 Oracle Solaris 信息库中的 *smb*。NTLMv2 是建议的最低安全性级别，可避免广为人知的一些安全漏洞。
- Preferred domain controller (首选域控制器) – 首选域控制器在加入“[Active Directory](#)” [228]域时使用。如果该控制器不可用，则 Active Directory 将根据 DNS SRV 记录和 Active Directory 站点来查找合适的域控制器。
- Active Directory site (Active Directory 站点) – 此站点在加入 Active Directory 域时使用。站点是计算机的逻辑集合，这些计算机全部通过高带宽、低延迟网络链路连接。如果配置了此属性，但未指定首选域控制器，则加入 Active Directory 域时将优先使用位于此站点内的域控制器，而不是外部的域控制器。
- Maximum # of server threads (最大服务器线程数) – 并发服务器线程（工作进程）的最大数目。默认值为 1024。
- Enable Dynamic DNS (启用动态 DNS) – 选择设备是否将使用动态 DNS 来更新 Active Directory 域中的 DNS 记录。默认值为关闭状态。
- Enable Oplocks (启用伺机锁定) – 选择设备是否将向 SMB 客户机授予伺机锁定。这将改进大多数客户机的性能。默认为打开状态。SMB 服务器向客户机进程授予伺机锁定以便客户机能够在该锁定就位时缓存数据。当服务器撤销伺机锁定时，客户机将其缓存的数据刷新到服务器。
- Restrict anonymous access to share list (限制匿名访问共享资源列表) – 如果启用了此选项，则客户机必须进行 SMB 服务验证才能接收共享资源列表。如果禁用，则匿名客户机可以访问共享资源列表。
- System Comment (系统注释) – 有意义的文本字符串。
- Idle Session Timeout (空闲会话超时) – 会话不活动状态的超时设置。
- Primary WINS server (主要 WINS 服务器) – 在 TCP/IP 设置中配置的主要 WINS 地址。
- Secondary WINS server (次要 WINS 服务器) – 在 TCP/IP 设置中配置的次要 WINS 地址。
- Excluded IP addresses from WINS (从 WINS 排除 IP 地址) – 从注册 WINS 排除 IP 地址。
- SMB Signing Enabled (启用 SMB 签名) – 使用 SMB 签名功能启用与 SMB 客户机的互操作性。如果某个包已签名，则将对签名进行验证。如果某个包未签名，则不进行签名验证即接受该包（如果不需要 SMB 签名 – 请参见下文）。

- **SMB Signing Required** (需要 SMB 签名) – 当需要 SMB 签名时，所有 SMB 包都必须签名，否则它们将被拒绝，并且不支持签名的客户机将无法连接到服务器。
- **Ignore zero VC** (忽略零 VC) – 当 SMB 客户机建立新的连接时，可能会请求设备通过将虚拟电路 (Virtual Circuit, VC) 数量指定为零，从客户机上清理所有以前的连接和文件锁定。但本协议对象不支持客户机或指定到相同主机的多个 DNS 条目的网络地址转换 (network address translation ,NAT)。在组合中，屏蔽或冗余网络位置之间的零 VC 请求可能会导致不相关的活动连接重置。默认情况下，零 VC 请求可防止过期文件锁定，但如果 SMB 会话在错误中断开，忽略零 VC 请求可解决该问题。

更改服务属性记录在[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)中。CLI 属性名称是上面列出的名称的短版本。

SMB 共享资源属性

当通过 SMB 导出共享资源时，必须以特定的方式设置几个[“共享资源属性” \[263\]](#)。

表 11-10 SMB 共享资源属性

属性	描述
第 12 章 共享资源、项目和模式	SMB 客户机要求不区分大小写，因此该属性必须为 "mixed" (混合) 或 "insensitive" (不区分大小写)。
第 12 章 共享资源、项目和模式	如果文件系统中允许使用非 UTF-8 文件名，则 SMB 客户机可能无法正常运行。
Non-Blocking Mandatory Locking (非阻塞强制性锁定)	要使字节范围锁定正常运行，必须启用此属性。
“共享资源协议” [290]	客户机用来引用共享资源的名称。有关该名称如何从 第 12 章 共享资源、项目和模式 继承的详细信息，请参见 “共享资源协议” [290] 文档。
“共享资源协议” [290]	一种 ACL，它在文件系统中存储的 ACL 之上又增加了一个访问控制层。有关此属性的详细信息，请参见 “共享资源协议” [290] 文档。

[第 12 章 共享资源、项目和模式](#)和[第 12 章 共享资源、项目和模式](#)属性只能在创建共享时设置。

NFS/SMB 互操作性

该设备支持可同时访问相同共享资源的[“NFS” \[185\]](#)和 SMB 客户机。要正确配置设备以实现 NFS/SMB 互操作性，必须配置以下组件：

- 配置[“Active Directory” \[228\]](#)服务。

- 制定“身份映射服务” [233]策略并配置服务。
- 配置 SMB。
- 在共享资源上配置访问控制、ACL 条目和 ACL 继承。

SMB 和 NFSv3 使用不同的访问控制模型。为获得最佳效果，请从 SMB 客户机的根目录上配置 ACL，因为 SMB 访问控制模型是更详细的模型。有关可继承的普通 ACL 条目的信息，请参见“Shares (共享资源) > Shares (共享资源) > Access (访问)” [297]文档。

SMB DFS 名称空间

分布式文件系统 (Distributed File System, DFS) 是通过 SMB 和 MSRPC 协议提供的一项虚拟技术。DFS 允许管理员对位于不同服务器上的共享文件夹进行分组，方法是透明地将它们连接到一个或多个 DFS 名称空间。DFS 名称空间是组织中的共享文件夹的虚拟视图。管理员可以选择要在名称空间中包含哪些共享文件夹，设计显示这些共享文件夹的分层结构，并确定在名称空间中显示的共享文件夹名称。当用户查看名称空间时，这些文件夹似乎驻留在单个高容量的文件系统中。用户可在名称空间中的文件夹之间导航，而无需知道承载数据的服务器名称或共享文件夹。

在每个系统上只能将一个共享资源置备为独立的 DFS 名称空间。不支持基于域的 DFS 名称空间。请注意，每个群集中只能置备一个 DFS 名称空间，即使每个群集节点都有一个单独的存储池也是如此。要将某个 SMB 共享资源置备为 DFS 名称空间，请使用 DFS 管理 MMC 管理单元创建一个独立的名称空间。

如果设备未加入“Active Directory” [228]域，则需要执行额外的配置，然后工作组用户才能修改 DFS 名称空间。要使 SMB 本地用户能够创建或删除 DFS 名称空间，该用户必须在服务器上创建一个单独的本地帐户。在下面的示例中，相关步骤使 SMB 本地用户 dfsadmin 能够操作 DFS 名称空间。

SMB Microsoft 独立 DFS 名称空间管理工具支持矩阵

下表列出了多个 Windows 操作系统版本上 Microsoft DFS 工具的操作（子命令/选项）。它标识设备上 DFS 服务支持的操作，以便管理设备上的独立 DFS 名称空间。

Microsoft Windows systems	XP 2003 2003 Vista 2008 2008 Win7						
			R2			R2	
	SP3	SP2	SP2	SP2	SP2	SP1	SP1
dfscmd CLI:							
/map [comment] [/restore]	y	y	y	y	y	y	y
/unmap	y	y	y	y	y	y	y
/add [/restore]	y	y	y	y	y	y	y
/remove	y	y	y	y	y	y	y
/view [/partial /full]	y	y	y	y	y	y	y

dfsutil CLI (old format):							
/addstdroot [/comment]	y	y	y	n	n	y	y
/remstdroot	y	y	y	n	n	y	y
/root:<DfsName> /view	n	n	n	y	y	y	y
/addlink [/comment]	NA	NA	NA	y	y	y	y
/removelink	NA	NA	NA	y	y	y	y
/state /display	NA	NA	NA	y	y	y	y
/state /enable	NA	NA	NA	y	y	y	y
/state /disable	NA	NA	NA	y	y	y	y
/ttl /display	NA	NA	NA	y	y	y	y
/ttl /set	NA	NA	NA	y	y	y	y
/server:<MachineName> /view	y	y	y	y	y	y	y
dfsutil CLI (new format):							
root addstd [comment]	NA	NA	NA	n	n	y	y
root remove	NA	NA	NA	n	n	y	y
root (view namespace)	NA	NA	NA	y	y	y	y
link add [comment]	NA	NA	NA	y	y	y	y
link remove	NA	NA	NA	y	y	y	y
link (view)	NA	NA	NA	y	y	y	y
target add	NA	NA	NA	y	y	y	y
target remove	NA	NA	NA	y	y	y	y
target (view)	NA	NA	NA	y	y	y	y
property comment (view)	NA	NA	NA	y	y	y	y
property comment set	NA	NA	NA	y	y	y	y
property ttl (view)	NA	NA	NA	y	y	y	y
property ttl set	NA	NA	NA	y	y	y	y
property state (view)	NA	NA	NA	y	y	y	y
property state offline	NA	NA	NA	y	y	y	y
property state online	NA	NA	NA	y	y	y	y
DFS GUI:							
add standalone root	y	y	y	n	n	n	n
remove standalone root	y	y	y	n	n	n	n
change root comment	y	y	y	n	n	n	n
change root timeout	y	y	y	n	n	n	n
add link	y	y	y	n	n	n	n
remove link	y	y	y	n	n	n	n
change link comment	y	y	y	n	n	n	n
change link timeout	y	y	y	n	n	n	n
add link's target	y	y	y	n	n	n	n
remove link's target	y	y	y	n	n	n	n
enable link's referral (target)	y	y	y	n	n	n	n
disable link's referral (target)	y	y	y	n	n	n	n
hide root	y	y	y	y	y	y	y
show root	y	y	y	y	y	y	y
display links	y	y	y	n	n	n	n
display targets	y	y	y	n	n	n	n
XP 2003 2003 Vista 2008 2008 Win7			R2			R2	
SP3 SP2 SP2 SP2 SP2 SP1 SP1							

注意：y – 支持 n – 不支持 NA – 不适用

- Solaris 不验证 DFS 链接目标。
- 用于修改及查看注释和超时 (TTL) 的 CLI 命令适用于 root 和链接。
- 用于查看状态的 CLI 命令适用于 root、root 目标、链接和链接目标。
- 用于修改状态的 CLI 命令适用于链接和链接目标。

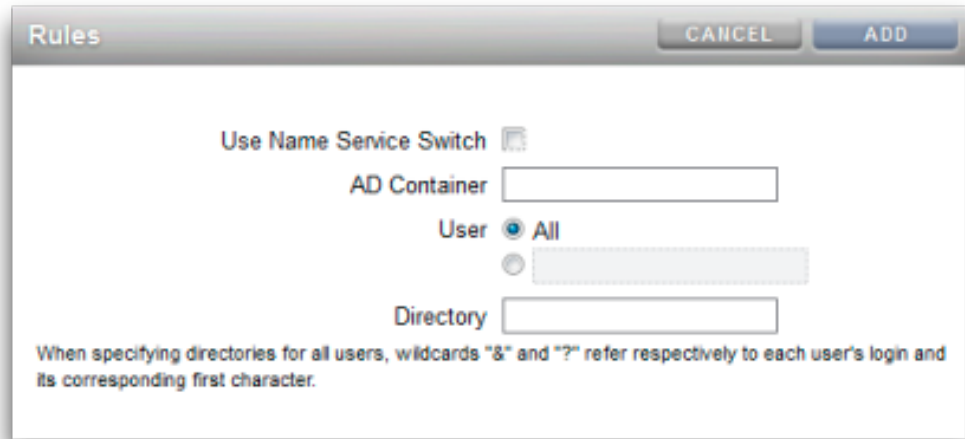
▼ 示例：操作 DFS 名称空间

1. 在服务器上为用户 `dfsadmin` 创建一个本地用户帐户。确保使用与首次在 Windows 计算机上创建本地用户时相同的密码。
2. 将 `dfsadmin` 添加到本地 SMB 组管理员。
3. 以 `dfsadmin` 用户身份在要从中修改 DFS 名称空间的 Windows 计算机上登录。

SMB 自动主目录服务

对于 Windows 文件共享，自动主目录可提供使用 SMB 协议访问文件系统的权限。对于通过 SMB 访问系统的用户，自动主目录可定义及维护主目录共享资源。自动主目录规则将 SMB 客户机映射到主目录。

图 11-2 设置自动主目录规则



- Name Service Switch (名称服务转换) – 打开或关闭名称服务转换 (Name Service Switch, NSS)。无法在创建 NSS 规则的同时为所有用户创建规则。
- AD Container (AD 容器) – 设置活动目录容器，例如：
dc=com，dc=fishworks，ou=Engineering，CN=myhome。
- User (用户) – 为所有用户或指定的用户设置自动主目录。当指定用户时，通配符 "&" 和 "?" 表示用户的登录及其对应的第一个字符。
- Directory (目录) – 设置规则的目录，例如：/export/wdp。

▼ 添加 SMB 自动主目录规则

1. 使用 **create** 命令添加自动主目录规则，使用 **list** 命令列出现有规则。此示例为用户 "Bill" 添加规则，然后列出规则：

```
twofish:> configuration services smb
twofish:configuration services smb> create
twofish:configuration services rule (uncommitted)> set use_nss=false
twofish:configuration services rule (uncommitted)> set user=Bill
twofish:configuration services rule (uncommitted)> set directory=/export/wdp
twofish:configuration services rule (uncommitted)> set container="dc=com,dc=fishworks,
ou=Engineering,CN=myhome"
twofish:configuration services rule (uncommitted)> commit
twofish:configuration services smb> list
RULE          NSS          USER          DIRECTORY          CONTAINER
```

```
rule-000    false    Bill           /export/wdp           dc=com,dc=fishworks,
ou=Engineering,CN=myhome
```

2. 可以使用通配符创建自主目录规则。& 字符匹配用户的用户名，? 字符匹配用户的用户名中第一个字符。下面的示例使用通配符来匹配所有用户：

```
twofish:configuration services smb> create
twofish:configuration services rule (uncommitted)> set use_nss=false
twofish:configuration services rule (uncommitted)> set user=*
twofish:configuration services rule (uncommitted)> set directory=/export/?/&
twofish:configuration services rule (uncommitted)> set container="dc=com,dc=fishworks,
ou=Engineering,CN=myhome"
twofish:configuration services rule (uncommitted)> commit
twofish:configuration services smb> list
RULE      NSS      USER      DIRECTORY      CONTAINER
rule-000   false    Bill      /export/wdp     dc=com,dc=fishworks,
ou=Engineering,CN=myhome
```

3. 还可以使用名称服务转换来创建自主目录规则：

```
twofish:configuration services smb> create
twofish:configuration services rule (uncommitted)> set use_nss=true
twofish:configuration services rule (uncommitted)> set container="dc=com,dc=fishworks,
ou=Engineering,CN=myhome"
twofish:configuration services rule (uncommitted)> commit
twofish:configuration services smb> list
RULE      NSS      USER      DIRECTORY      CONTAINER
rule-000   true     ou=Engineering,CN=myhome  dc=com,dc=fishworks,
```

SMB 本地组

本地组是由域用户构成的组，它向这些用户授予附加特权。

表 11-11 SMB 本地组

组	说明
Administrators (管理员)	管理员可以绕过文件权限来更改文件的所有权。
Backup Operators (备份操作员)	备份操作员可以绕过文件访问控制来备份和恢复文件。

▼ 将用户添加到 SMB 本地组

- 要添加用户，请执行以下操作：

```
twofish:configuration services smb> groups
twofish:configuration services smb groups> create
twofish:configuration services smb member (uncommitted)> set user=Bill
twofish:configuration services smb member (uncommitted)> set group="Backup Operators"
twofish:configuration services smb member (uncommitted)> commit
twofish:configuration services smb groups> list
MEMBER      USER      GROUP
member-000  WINDOMAIN\Bill      Backup Operators
```

SMB 本地帐户

本地帐户和用户 ID 映射到 Windows 用户 ID。请注意，*guest* 帐户是一个特殊的只读帐户，无法配置为用于在设备中读取/写入。

SMB MMC 集成

Microsoft 管理控制台 (Microsoft Management Console, MMC) 是包含已注册组件（称为管理单元）的一个可扩展框架，它可针对网络上的本地系统和远程系统提供全面的管理功能。"Computer Management"（计算机管理）是 Microsoft 管理控制台工具的集合，这些工具可用于配置、监视和管理本地和远程服务与资源。

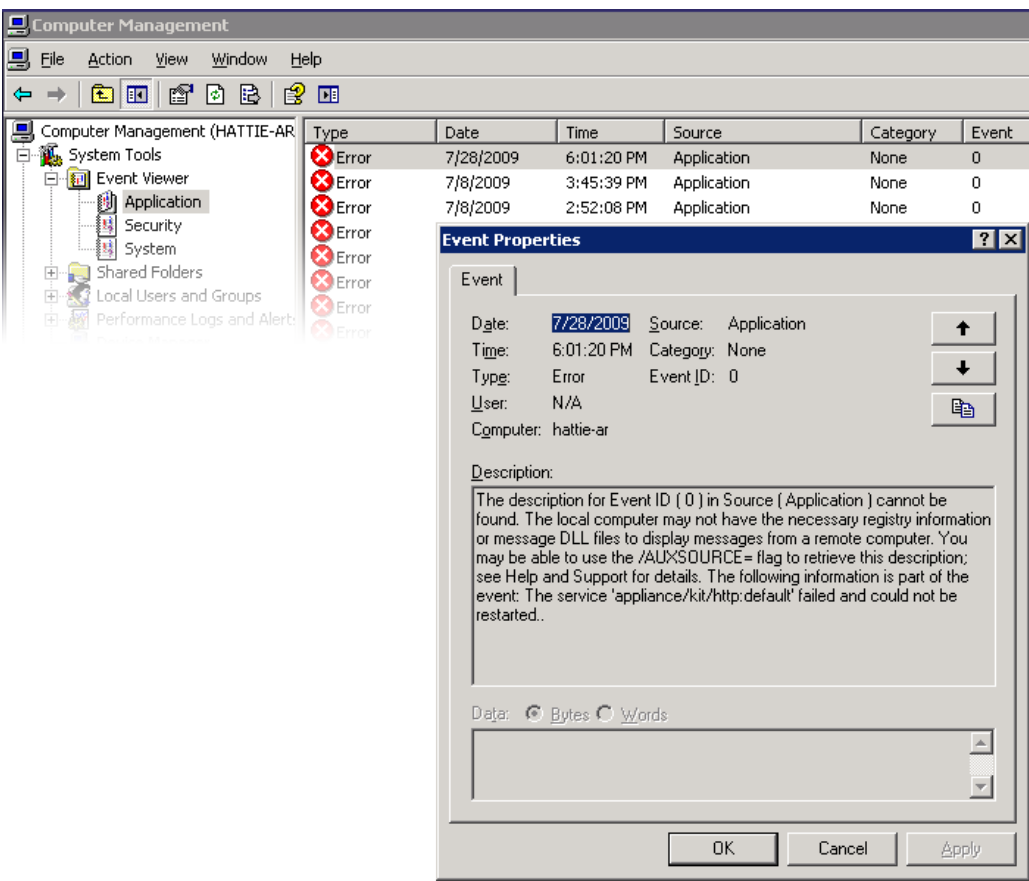
要在 Sun ZFS Storage 7000 设备上以工作组模式使用 MMC 功能，请确保将要使用管理控制台的 Windows 管理员添加到设备上的管理员"["local group"（本地组）](#)" [191]。否则，在尝试使用 MMC 连接到设备时，在管理客户机上可能会收到 Access is denied 或类似错误。

Sun ZFS Storage 7000 设备支持以下计算机管理工具：

SMB 事件查看器

使用事件查看器 MMC 管理单元，可以显示应用程序日志、安全日志和系统日志。这些日志显示 Sun ZFS Storage 7000 系统的警报、审计和系统日志的内容。下面的屏幕截图显示了一个错误事件的应用程序日志和属性对话框。

图 11-3 SMB 事件查看器



SMB 共享资源管理

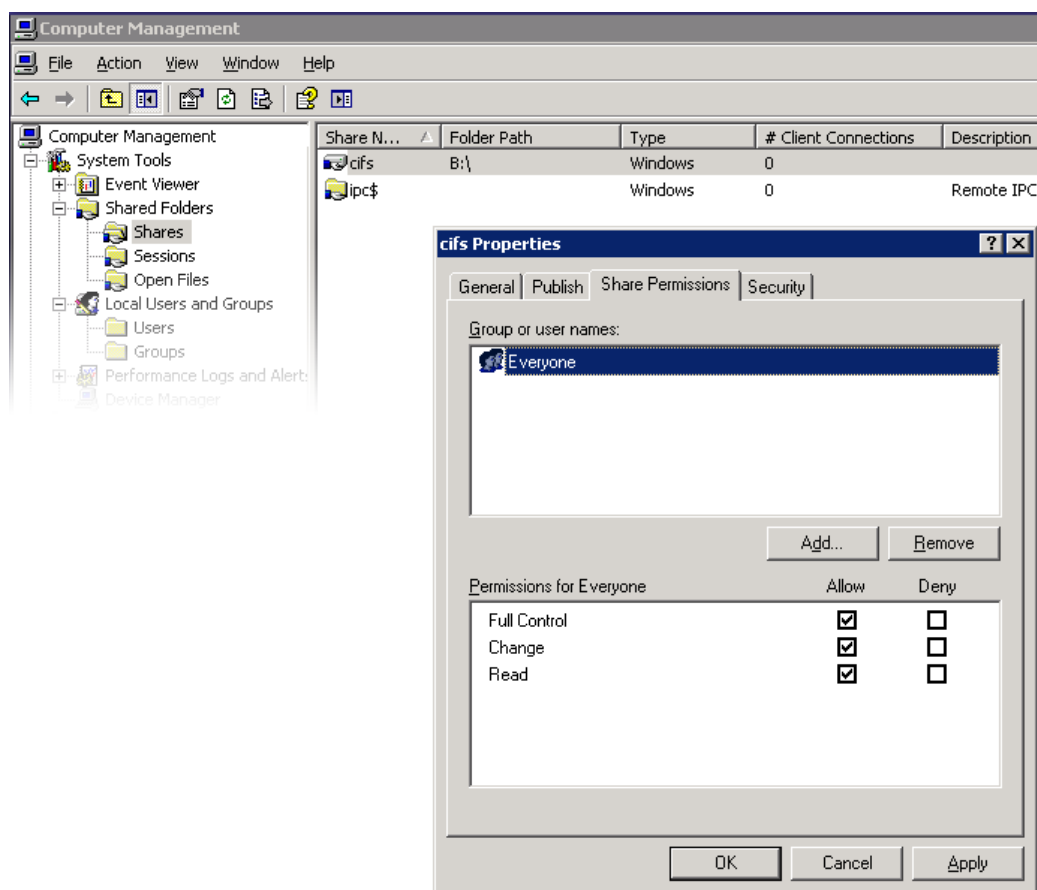
对共享资源管理的支持包括：

- 列出共享资源
- 在共享资源上设置 ACL
- 更改共享权限
- 设置共享资源的说明

当前未通过 MMC 支持的功能包括：

- 添加或删除共享资源
- 设置客户端缓存属性
- 设置允许的最大用户数属性

图 11-4 SMB 共享资源权限属性



SMB 用户、组和连接

支持以下功能：

- 查看本地 SMB 用户和组
- 列出用户连接，包括列出每个连接上打开的文件数

- 关闭用户连接
- 列出打开的文件，包括列出文件上的锁数和文件打开模式
- 关闭打开的文件

图 11-5 每个连接上打开的文件

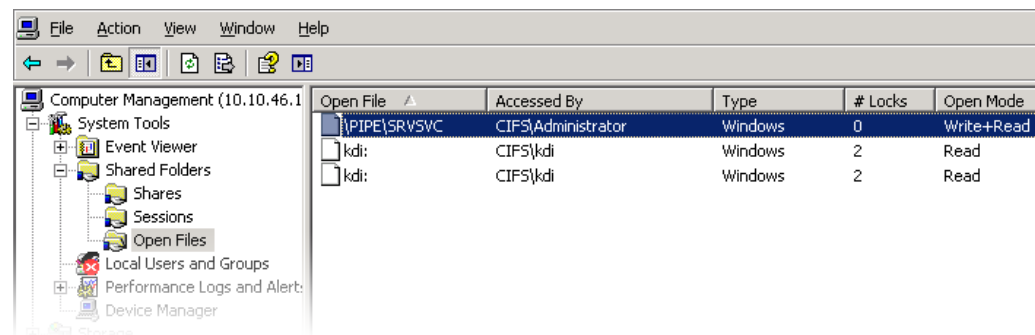
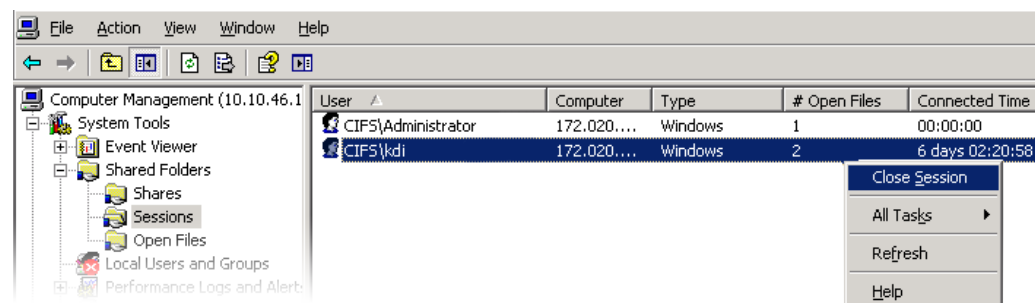


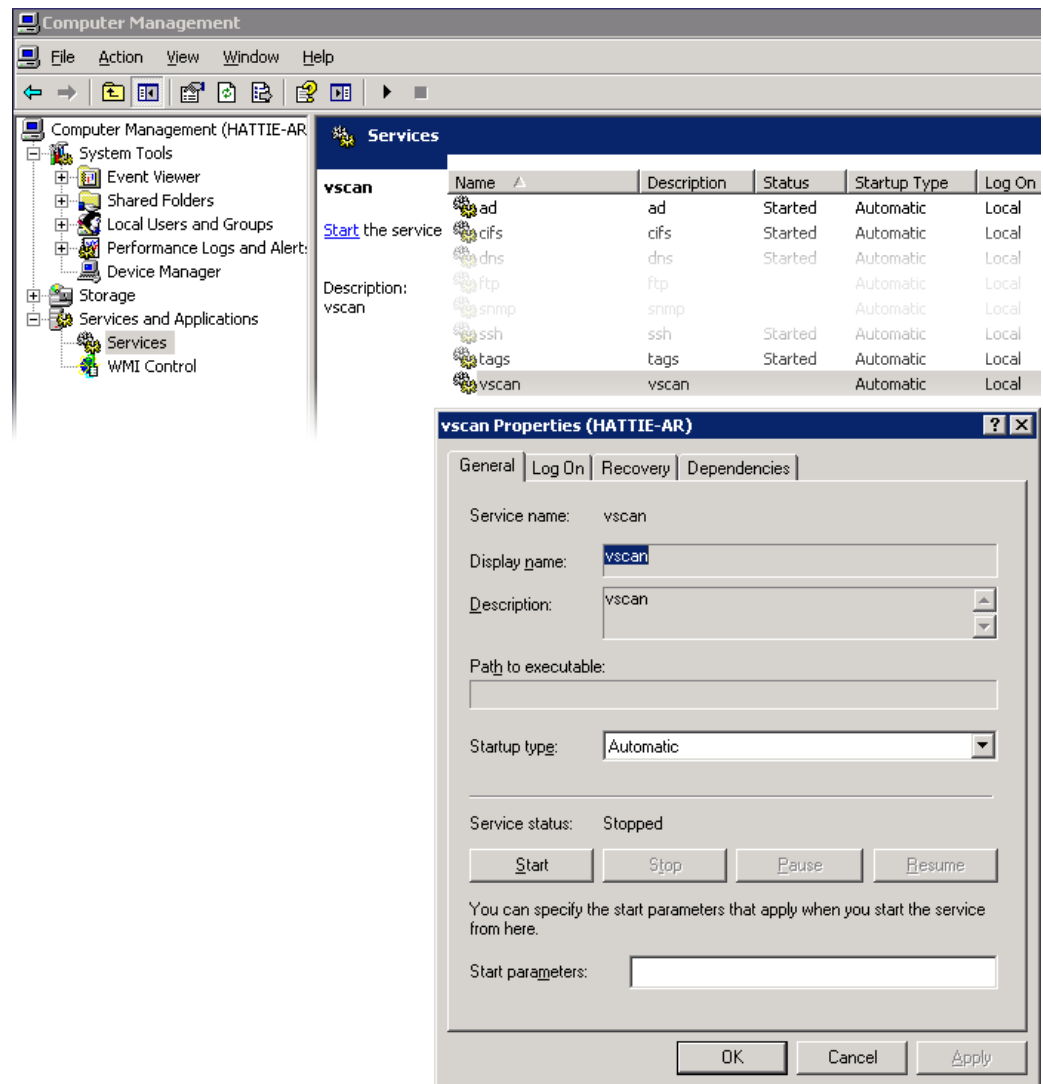
图 11-6 打开的会话



列出 SMB 服务

支持包括列出 ZFSSA 服务。无法使用计算机管理 MMC 应用程序启用或禁用服务。下面的屏幕抓图显示了 vsan 服务的常规属性。

图 11-7 vscan 属性



为确保只有适当的用户有权执行管理操作，对于使用 MMC 远程执行的操作施加了一些访问限制。

表 11-12 用户及允许的操作

用户	允许的操作
一般用户	列出共享资源。
"Administrators" 或 "Power Users" 组的成员	管理共享资源，列出用户连接。
"Administrators" 组的成员	列出打开的文件以及关闭文件、断开用户连接、查看服务和事件日志。

使用 BUI 配置 SMB

▼ 初始配置

可以使用 BUI 或 CLI 完成设备的初始配置，并且花费的时间应当少于 20 分钟。使用 BUI 或 CLI 的 "Maintenance"（维护）> "System"（系统）上下文以后还可以再次执行初始设置。通常，初始配置将带您完成以下 BUI 步骤。

1. 配置网络设备、数据链路和接口。
2. 使用 "Datalink"（数据链路）添加或 "Interface"（接口） 图标，或通过将设备拖放到数据链路或接口列表来创建接口。
3. 设置所需的属性并单击 "Apply"（应用）按钮将其添加到列表中。
4. 根据需要每个接口设置为活动或备用。
5. 单击页面顶部的 "Apply"（应用）按钮提交您的更改。
6. 配置 DNS。
7. 提供基本域名。
8. 提供至少一台服务器的 IP 地址，该服务器必须能够解析域名称空间的 Active Directory 部分中的主机名和服务器记录。
9. 配置 NTP 验证密钥以确保时钟同步。
10. 单击  图标添加新密钥。
11. 为新密钥指定编号、类型和私钥值并应用更改。该密钥将在每个指定的 NTP 服务器旁显示为选项。

12. 将该密钥与适当的 NTP 服务器关联并应用更改。为确保时钟同步，请将设备和 SMB 客户机配置为使用相同的 NTP 服务器。
13. 将 Active Directory 指定为用户和组的目录服务。
14. 设置目录域。
15. 单击 "Apply" (应用) 按钮提交您的更改。
16. 配置存储池。
17. 单击  图标添加新池。
18. 设置池名称。
19. 在 "Allocate and verify storage" (分配和验证存储) 屏幕上，为存储池配置 JBOD 分配。JBOD 分配可以为 "none" (无)、"half" (半) 或 "all" (全部)。如果未检测到 JBOD，请检查 JBOD 布线和电源。
20. 单击 "Commit" (提交) 按钮前进到下一屏幕。
21. 在 "Configure Added Storage" (配置添加的存储) 屏幕上，选择所需的数据配置文件。每项都在可用性、性能和容量方面进行了评级。请使用这些评级确定适合您的业务需求的最佳配置。
22. 单击 "Commit" (提交) 按钮以激活配置。
23. 配置远程支持。
24. 如果设备没有直接连接到 Internet，请配置 Oracle 用来与远程支持服务进行通信的 HTTP 代理。
25. 输入您的联机帐户用户名和密码。将显示一份隐私声明供您阅读。
26. 选择要向其注册的清单组。每个帐户的默认组与帐户用户名相同，以 '\$' 为前缀。
27. 提交您的初始配置更改。

▼ Active Directory 配置

1. 在 Active Directory 域中为设备创建一个帐户。有关详细说明，请参阅 Active Directory 文档。

2. 在 "Configuration" (配置) > "Services" (服务) > "Active Directory" 屏幕上, 单击 "Join Domain" (加入域) 按钮。
3. 指定 Active Directory 域、管理用户、管理密码, 然后单击 "Apply" (应用) 按钮提交更改。

▼ 项目和共享资源配置

1. 创建项目。
2. 在 "Shares" (共享资源) 屏幕上, 单击  图标以展开 "Projects" (项目) 面板。
3. 单击 "Add..." (添加...) 链接以添加新项目。
4. 指定项目名称并应用更改。
5. 从 "Projects" (项目) 面板中选择新项目。
6. 单击  图标添加新文件系统。
7. 单击该文件系统的  图标。
8. 单击 "General" (常规) 链接并取消选中 "Inherit from project" (从项目继承) 复选框。
9. 选择 /export 下的挂载点, 即使 SMB 共享资源是按资源名称进行访问的。
10. 在项目的 "Protocols" (协议) 屏幕上, 将资源名称设置为 "on"。
11. 为项目启用 sharesmb 和共享资源级 ACL。
12. 单击 "Apply" (应用) 按钮以激活配置。

▼ SMB 数据服务配置

1. 在 "Configuration" (配置) > "Services" (服务) > "SMB" 屏幕上, 单击  图标启用服务。
2. 根据本页面中 "属性" 部分中的建议设置 SMB 属性, 然后单击 "Apply" (应用) 按钮以激活配置。

3. 在 "Configuration" (配置) > "Services" (服务) > "SMB" 屏幕上, 单击 "Autohome" (自动主目录) 链接, 以根据上面“自动主目录规则”部分中的说明设置自动主目录规则, 来将 SMB 客户机映射到主目录, 然后单击 "Apply" (应用) 按钮激活配置。
4. 在 "Configuration" (配置) > "Services" (服务) > "SMB" 屏幕上单击 "Local Groups" (本地组) 链接, 并根据上面 "Local Groups" (本地组) 部分中的说明使用  图标将管理员或备份操作员用户添加到本地组, 然后单击 "Apply" (应用) 按钮激活配置。

FTP 服务

通过 FTP (File Transfer Protocol, 文件传输协议) 服务, 可以从 FTP 客户机访问文件系统。不允许匿名登录, 用户必须使用服务中配置的名称服务进行验证。

FTP 属性

FTP 常规设置

表 11-13 FTP 常规设置

属性	说明
Port (端口) (用于传入连接)	FTP 侦听的端口。默认值是 21
Maximum # of connections (最大连接数) ("0" 表示没有限制)	这是最大并发 FTP 连接数。可将此值设置为包含预期的并发用户数。默认情况下, 此值为 30, 因为每个连接都会创建一个系统进程, 因此允许太多连接 (数千) 可能会构成 DoS 攻击
Turn on delay engine to prevent timing attacks (启用延迟引擎以防止计时攻击)	这将在验证过程中插入短暂的延迟以欺骗尝试通过计时手段猜测用户名的攻击者。启用此设置将提高安全性
Default login root (默认 root 登录)	FTP 登录位置。默认值为 "/", 指向共享资源分层结构的顶层。所有用户在成功进行 FTP 服务验证后都将登录到此位置
Logging level (事件记录级别)	proftpd 日志的详细级别。
Permissions to mask from newly created files and dirs (从新创建的文件和目录屏蔽的权限)	创建文件时要删除的文件权限。默认情况下, 会屏蔽组和全局写入以防止每个人都能对最新的上载执行写操作

FTP 安全设置

表 11-14 FTP 安全设置

属性	说明
Enable SSL/TLS (启用 SSL/TLS)	允许经过 SSL/TLS 加密的 FTP 连接。这将确保加密 FTP 事务。默认设置为禁用。
Port for incoming SSL/TLS connections (端口 (用于传入的 SSL/TLS 连接))	经过 SSL/TLS 加密的 FTP 服务侦听的端口。默认值为 21。
Permit root login (允许 root 用户登录)	允许 root 用户的 FTP 登录。默认情况下将其设置为禁用，因为 FTP 验证是纯文本，会带来网络嗅探攻击的安全风险
Maximum # of allowable login attempts (允许的最大登录尝试次数)	FTP 连接断开之前的失败登录尝试次数；此后用户必须重新连接才能再次尝试。默认值为 3
Permit foreign data connection addresses (允许外来数据连接地址)	允许外来 FTP 连接以启用 FTP 服务器之间文件的直接传输。默认情况下，此属性处于禁用状态。

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

FTP 日志

表 11-15 FTP 日志

日志	说明
proftpd	记录 FTP 事件，包括成功登录和不成功登录尝试
proftpd_xfer	文件传输日志
proftpd_tls	记录与 SSL/TLS 加密相关的 FTP 事件

使用 BUI 配置 FTP

▼ 允许对共享资源进行 FTP 访问

1. 转到 "Configuration" (配置) -> "Services" (服务)
2. 确保 FTP 服务已启用且处于联机状态。如果不是，请启用该服务。

3. 在 "Shares"（共享资源）屏幕中选择或添加一个共享资源。
4. 转到 "Protocols"（协议）部分，检查是否启用了 FTP 访问。此处还可以设置访问模式（读取/读取+写入）。

HTTP 服务

通过 HTTP 服务，可以使用 HTTP 和 HTTPS 协议以及 HTTP 扩展 WebDAV（Web based Distributed Authoring and Versioning，Web 分布式创作和版本控制）访问文件系统。这样一来，客户机便可通过 Web 浏览器访问共享文件系统；如果客户机软件支持，还可将其当作本地文件系统进行访问。用于访问这些 HTTP 和 HTTPS 共享资源的 URL 分别采用以下格式：

http://hostname/shares/mountpoint/share_name

https://hostname/shares/mountpoint/share_name

HTTPS 服务器使用自签名安全证书。

HTTP 属性

表 11-16 HTTP 属性

属性	说明
Require client login（需要客户机登录）	客户机必须先进行验证才能访问共享资源，客户机对其创建的文件具有所有权。如果不设置此属性，则创建的文件将由 HTTP 服务用户 "nobody" 拥有。请参见下文有关验证的部分。
Protocols（协议）	选择支持 HTTP 和/或 HTTPS 的访问方法。
HTTP Port（HTTP 端口）（用于传入连接）	HTTP 端口，默认值为 80
HTTPS Port (for incoming secure connections)（HTTPS 端口（用于传入的安全连接））	HTTP 端口，默认值为 443

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

HTTP 验证和访问控制

如果启用了 "Require client login"（需要客户机登录）选项，设备将拒绝访问不提供本地用户、NIS 用户或 LDAP 用户的有效验证凭证的客户机。不支持 Active Directory 验证。

仅支持基本的 HTTP 验证。请注意，除非使用的是 HTTPS，否则将传输未加密的用户名和密码，这可能并不适合所有环境。

通常，经过验证的用户拥有的 HTTP 权限与其拥有的 NFS 或 FTP 权限相同。经过验证的用户将拥有其创建的文件和目录，和其他协议的对待方式一样。为了进行访问控制，将特权用户（uid 小于 100 的用户）视为 "nobody"。特权用户创建的文件归 "nobody" 所有。

如果禁用 "Require client login"（需要客户机登录）选项，设备不会尝试验证客户机（即使客户机提供了凭证）。为了进行访问控制，新创建的文件归 "nobody" 所有，并将所有用户都视为 "nobody"。

不管是否进行验证，都不会屏蔽已创建文件和目录的任何权限。创建的文件具有 Unix 权限 666（每个人都可读、可写），创建的目录具有 Unix 权限 777（每个人都可读、可写且可执行）。

HTTP 日志

表 11-17 HTTP 日志

日志	说明
network-http:apache22	HTTP 服务日志

配置 HTTP

▼ 允许对共享资源进行 HTTP 访问

1. 转到 "Configuration"（配置）-> "Services"（服务）
2. 检查 HTTP 服务是否已启用且处于联机状态。如果不是，请启用该服务。
3. 在 "Shares"（共享资源）屏幕中选择或添加一个共享资源。
4. 转到 "Protocols"（协议）部分，检查是否启用了 HTTP 访问。此处还可以设置访问模式（读取/读取+写入）。

NDMP 服务

通过 NDMP（Network Data Management Protocol，网络数据管理协议）服务，系统可以参与由远程 NDMP 客户机（称为数据管理应用程序 (Data Management Application,

DMA)) 控制的基于 NDMP 的备份和恢复操作。通过使用 NDMP，可将设备用户数据（即存储在设备上管理员创建的共享资源中的数据）备份和恢复到本地连接的磁带设备和远程系统。本地连接的磁带设备还可向 DMA 显示以进行远程系统的备份和恢复。

NDMP 不能用于备份和恢复系统配置数据。而是使用 `[[Maintenance:System:ConfigurationBackup]配置备份和恢复]]` 功能。

NDMP 本地配置与远程配置

设备支持使用本地配置（磁带机以物理方式连接到设备）和远程配置（数据传输到同一个网络上的其他系统）进行备份和恢复。在这两种配置中，备份都必须通过支持的 DMA 管理。

在本地配置中，支持的磁带设备（包括磁带机和转换器（机械手））使用在启动器模式下配置的受支持的 SCSI 卡或光纤通道 (Fibre Channel, FC) 卡以物理方式连接到系统。可在“[NDMP 状态](#)” [\[57\]](#) 屏幕上查看这些设备。NDMP 服务在 DMA 扫描设备时，会将这些设备呈现给 DMA。在 DMA 中配置了这些设备后，即可将其用于备份和恢复同一个网络上的设备或其他系统。在系统中添加或移除了磁带机或转换器之后，可能需要重新引导之后，NDMP 服务才可识别更改。此后，由于磁带机名称可能已发生更改，因此可能需要重新配置 DMA。

在远程配置中，磁带设备未以物理方式连接到要备份和恢复的系统（数据服务器），而是连接到运行 DMA 的系统或单独的系统（磁带服务器）。这些通常称为“3 向配置”，因为 DMA 控制其他两个系统。在这些配置中，数据流通过 IP 网络在数据服务器和磁带服务器之间传输。

NDMP 备份格式和类型

NDMP 协议未指定备份数据格式。设备支持三种备份类型，它们对应于不同的实施和磁带存储格式。DMAs 可以使用以下 NDMP 环境变量 “TYPE”（类型）值选择备份类型：

表 11-18 NDMP 备份格式和类型

备份类型	详细信息
dump	基于文件，仅适用于文件系统。支持文件历史记录和直接访问恢复 (Direct Access Recovery, DAR)。
tar	基于文件，仅适用于文件系统。支持文件历史记录和直接访问恢复 (Direct Access Recovery, DAR)。
zfs	基于共享资源，适用于文件系统和卷。不支持文件历史记录或直接访问恢复 (Direct Access Recovery, DAR)，但备份某些数据集的速度可能更快。仅支持使用 NDMPv4。

没有标准的 NDMP 数据流格式，因此只能在运行兼容软件的 7000 系列设备上恢复在设备上生成的备份流。将来版本的设备软件一般可以恢复通过较早版本的软件备份的数据流，反之则不一定成立。例如，“zfs”备份类型是 2010.Q3 中的新类型，运行 2010.Q1 或更早版本的系统无法恢复在 2010.Q3 下使用“zfs”类型创建的备份流。

使用 "dump" 和 "tar" 进行 NDMP 备份

使用 "dump" 和 "tar" 备份类型备份时，管理员通过文件系统路径（称为备份路径）指定要备份的数据。例如，如果管理员配置了 `/export/home` 的备份，则将备份在该路径上挂载的共享资源。同样，如果将备份流恢复到 `/export/code`，则这就是恢复文件的路径，即使文件是从其他路径备份时也是如此。

只能将现有共享资源的挂载点路径或者现有共享资源中包含的路径指定为备份路径。如果备份路径与某个共享资源的挂载点相同，则仅备份该共享资源。否则，路径必须包含在某个共享资源中，此时将仅备份该共享资源在此路径下的部分。在两种情况下，都不备份在备份路径下指定的共享资源内部挂载的其他共享资源；必须单独指定这些共享资源进行备份。

Snapshots（快照） – 如果备份路径指定了一个活动文件系统（例如 `/export/code`）或者活动文件系统中包含的路径（例如 `/export/code/src`），设备将立即创建新的快照并从该快照备份指定的路径。备份完成时，将销毁该快照。如果备份路径指定了一个快照（例如 `/export/code/.zfs/snapshot/mysnap`），则不会创建新快照，系统将从指定的快照进行备份。

Share metadata（共享资源元数据） – 为了简化复杂共享资源配置的备份和恢复，“dump”和“tar”备份包括了与备份路径关联的项目和共享资源的共享资源元数据。此元数据说明设备上的共享资源配置，包括协议共享属性、配额属性以及在“Shares”（共享资源）屏幕上配置的其他属性。不要将其与也是使用 NDMP 进行备份和恢复的目录结构和文件权限之类的文件系统元数据混淆。

例如，如果备份 `/export/proj`，则将备份挂载点以 `/export/proj` 开头的所有共享资源的共享资源元数据，以及这些共享资源的父项目的共享资源元数据。同样，如果备份 `/export/someshare/somedir`，并且某个共享资源在 `/export/someshare` 处挂载，则将备份该共享资源及其项目的共享资源元数据。

恢复时，如果恢复目标路径不在现有共享资源中，则将根据需要，使用备份中存储的项目和共享资源的原始属性重新创建备份流中的项目和共享资源。例如，如果备份 `/export/foo`（其中包含项目 `proj1` 以及共享资源 `share1` 和 `share2`），然后销毁项目并从备份恢复，则在恢复操作过程中将使用这两个共享资源和该项目的已备份属性重新创建它们。

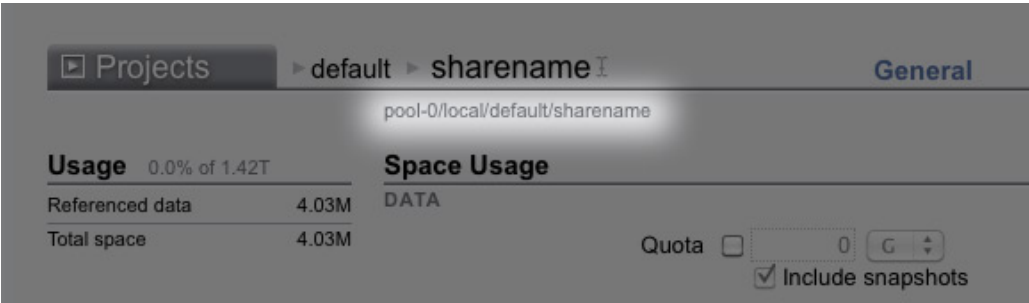
在恢复过程中，如果已经存在要自动重新创建的项目，则将使用现有项目，而不会自动创建新项目。如果存在要自动重新创建的共享资源，并且其挂载点与设备根据原始备份路径和恢复目标预期的挂载点匹配，则使用现有的共享资源，而不会自动创建新共享资源。否则，将根据备份中的元数据自动创建一个新共享资源。如果已经存在同名（但挂载点不同）的共享资源，则将为新创建的共享资源指定一个以“ndmp-”开头的唯一名称并指定正确的挂载点。

建议恢复设备上已不再有其数据集的备份流，使设备可以重新创建备份流中指定的数据集，或者预创建恢复的目标共享资源。这两种做法都可以避免出现与上述自动共享资源创建相关的意外结果。

使用 "zfs" 进行 NDMP 备份

使用 "zfs" 类型备份时，管理员按数据在设备上的规范名称指定要备份的数据。可在 BUI 中共享资源名称下方找到此名称：

图 11-8 NDMP 共享资源名称



也可在 CLI 中找到以 canonical_name 属性值表示的此名称。规范名称不以前导 '/' 开头，但在配置备份路径时，规范名称必须加前缀 '/'。

可以为使用 "zfs" 类型的备份指定项目和共享资源。如果按原样指定规范名称，则将创建一个新快照并将其用于备份。可以使用 '@snapshot' 后缀为备份指定特定的快照；在这种情况下，不会创建新快照，将备份指定的快照。例如：

表 11-19 规范名称和共享资源备份

规范名称	备份的共享资源
pool-0/local/default	名为 "default" 的本地项目及其所有共享资源的新快照。
pool-0/local/default@yesterday	本地项目 "default" 的命名快照 "yesterday" 及该项目具有快照 "yesterday" 的所有共享资源。
pool-0/local/default/code	本地项目 "default" 中的共享资源 "code" 的新快照。"code" 可以是文件系统或卷。
pool-0/local/default/code@yesterday	本地项目 "default" 中的共享资源 "code" 的命名快照。"code" 可以是文件系统或卷。

因为使用 "zfs" 备份类型的基于级别的增量备份需要上一个增量的基础快照，因此为其创建新快照的级别备份的默认行为是保留新快照，从而使其可用于后续的增量备份。如果 DMA 通过设置 UPDATE=n 指示备份不用于后续增量备份，则将在备份之后销毁新创建的快照。在任何情况下都不会在备份后销毁现有的用户快照。有关详细信息，请参见下文的“增量备份”。

Share metadata（共享资源元数据）– 共享资源元数据（即共享资源配置）始终包括在 "zfs" 备份中。在恢复使用 "zfs" 类型的完整备份时，目标项目或共享资源必须尚不存在。将根据备份流中的元数据重新创建目标项目或共享资源。在恢复使用 "zfs" 类型的增量备份时，必须已经存在目标项目或共享资源。将根据备份流中的元数据更新目标项目或共享资源的属性。有关详细信息，请参见下文的“增量备份”。

NDMP 增量备份

设备支持以上所有备份类型的基于级别的增量备份。为指定级别备份，DMA 通常会指定以下三个环境变量：

变量	详细信息
LEVEL	表示备份级别的整数 0 到 9。
DMP_NAME	指定特定的增量备份集。通过指定不同的 DMP_NAME 值，可以同时使用多个级别增量备份集。
UPDATE	指示此备份是否可用作后续增量备份的基础

按定义，级别 N 备份包括使用的 LEVEL 低于 N 的同一个共享资源的同一个备份集（通过 "DMP_NAME" 指定）上次备份以来发生更改的所有文件。级别 0 备份始终包括所有文件。如果 UPDATE 的值为 "y"（默认值），则将记录当前备份，以便级别大于 N 的未来备份使用此备份作为基础。这些变量通常由 DMA 管理，而不需要由管理员直接配置。

以下是一个增量备份调度表样例：

表 11-20 示例增量备份调度表

日期	详细信息
每月第一天	级别 0 备份。备份包含共享资源中的所有文件。
每月 7 日、14 日和 21 日	级别 1 备份。备份包含自上次完整（每月）备份以来发生更改的所有文件
每天	级别 -2 备份。备份包含自上次级别 1 备份以来发生更改的所有文件

要恢复文件系统在 24 日的状态，管理员通常将 1 日的级别 0 备份恢复到新的共享资源，然后恢复 21 日的级别 1 备份，再恢复 24 日的级别 2 备份。

要实施基于级别的增量备份，设备必须跟踪每个共享资源的级别备份历史记录。对于 "tar" 和 "dump" 备份，级别备份历史记录保留在共享资源元数据中。增量备份会遍历文件系统，并将上一级别备份以来经过修改的文件都包括在内。在恢复时，系统只需恢复备份流中的所有文件。因此，在上面的示例中，可以将 24 日的级别 2 备份恢复到任何文件系统上；即使目标文件系统与文件备份时的文件系统不匹配，也可以恢复该备份流中包含的文件。但是，最佳做法是使用上面所述的过程，从空树开始恢复之前级别的备份来恢复原始文件系统状态。

为了实施高效的基于级别的 "zfs" 类型增量备份，系统将使用一种不同的方法。增量集中的备份不会销毁用于备份的快照，而是将其留在系统上。后续的增量备份使用此快照作为基础，快速确定发生更改的文件系统块，并生成备份流。因此，如果要创建后续的增量备份，就不能销毁 NDMP 服务在备份之后保留的快照。

此行为的另一个重要后果是：为了恢复增量流，文件系统状态必须与其在增量流的基础快照上的状态完全匹配。换言之，要恢复级别 2 备份，文件系统必须与上一个级别 1 备份完成时完全相同。请注意，上述常用过程可保证这一点，因为恢复 24 日的级别 2 备份流时，系统与其在 21 日的级别 1 备份完成时完全相同（因为刚恢复该备份）。

如果试图将增量 "zfs" 备份流恢复到的文件系统的最新快照与该增量流的基础快照不匹配，或者文件系统自该快照之后发生了更改，NDMP 服务将报告错误。可以配置 NDMP 服务以回滚到恢复刚开始之前的基础快照，方法是：指定 NDMP 环境变量 "ZFS_FORCE" 的值 "y"，或者配置 NDMP 服务的 "Rollback datasets"（回滚数据集）属性（请参见下文的“属性”）。

NDMP 属性

NDMP 服务配置包含以下属性：

表 11-21 NDMP 属性

属性	说明
Version（版本）	DMA 支持的 NDMP 版本
TCP port (v4 only)（TCP 端口（仅 v4））	NDMP 默认连接端口为 10000。NDMPv3 始终使用此端口。NDMPv4 允许在需要时使用其他端口。
Default restore pool(s)（默认恢复池）	在进行使用 "tar" 或 "dump" 的完整恢复时，如果目标上未挂载共享资源，系统将重新创建数据集。由于 NDMP 协议仅指定挂载点，因此系统将选择要在其中重新创建项目和共享资源的池。在具有多个池的系统上，此属性允许您指定一个或多个池。仅在每个机头上都有活动池的群集中才需要指定多个池。您必须确保此列表与任何存储配置更改保持同步。如果不

属性	说明
	存在任何池或没有池联机，系统将随机选择一个默认池。
Ignore metadata-only changes (忽略纯元数据更改)	指示系统仅备份内容发生了更改的文件，忽略仅元数据（如权限或所有权）发生了更改的文件。此选项仅适用于增量 "tar" 和 "dump" 备份，默认情况下处于禁用状态。
Allow token-based backup (允许基于标记进行备份)	对于 ZFS 备份启用或禁用基于标记的方法。默认情况下，此属性处于禁用状态。
ZFS rollback before restore (v4 only) (恢复之前进行 ZFS 回滚 (仅 v4))	仅适用于 "zfs" 类型的备份。确定在恢复增量备份时系统是否将目标项目和共享资源回滚到用作增量恢复基础的快照。如果回滚项目和共享资源，将丢失自该快照以来所做的所有更改。此设置通常由 DMA 通过 "ZFS_FORCE" 环境变量控制（请参见上文中的“增量备份”），但可将此属性用于覆盖 DMA 设置以便始终回滚或从不回滚这些数据集。除非已经手动回滚了这些数据集，否则不将其回滚将导致恢复失败。此属性适用于不允许管理员配置定制环境变量（如 ZFS_FORCE）的 DMA。
Allow direct access recovery (允许直接访问恢复)	支持系统在恢复操作过程中通过位置（而不是顺序搜索）查找文件。启用此选项可减少从许多磁带恢复少量文件所用的时间。若要以后能够恢复单个文件，必须在备份时指定此选项。
Restore absolute paths (v3 only) (恢复绝对路径 (仅 v3))	指定在恢复文件时，同时恢复该文件的完整绝对路径（而不只是文件本身）。默认情况下此选项处于禁用状态。
DMA tape mode (for locally attached drives) (DMA 磁带模式 (适用于本地连接的驱动器))	指定 DMA 是预期 System V 还是 BSD 语义。默认值为 System V，这是适用于大多数 DMA 的推荐设置。此选项仅适用于通过 NDMP 导出的本地连接的磁带机。请参考 DMA 文档以了解您的 DMA 期望的模式。更改此选项仅会更改 DMA 扫描设备时导出的设备，因此，在更改此设置后，需要在 DMA 中重新配置磁带设备。
DMA username and password (DMA 用户名和密码)	用于验证 DMA。系统使用 MD5 进行用户验证

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

NDMP 日志

表 11-22 NDMP 日志

日志	说明
system-ndmpd:default	NDMP 服务日志

远程复制

远程复制服务可便于将项目和共享资源复制到其他 Oracle ZFS Storage Appliance 以及从这些设备复制项目和共享资源。[第 13 章 复制](#)文档中将详细介绍该功能。

启用此服务时，设备将从其他设备接收复制更新以及发送本地项目和共享资源的复制更新（根据为其配置的操作）。禁用此服务时，传入的复制更新将会失败，且不会复制任何本地项目和共享资源。

此服务没有任何属性，但可允许管理员查看将数据复制到此设备的设备（在 "Sources"（源）下），以及配置可将此设备的数据复制到其中的设备（在 "Targets"（目标）下）。有关管理远程复制的详细信息，可在[第 13 章 复制](#)文档中找到。

影子迁移

影子迁移服务允许从外部源或内部源自动迁移数据。[第 14 章 影子迁移](#)中将详细介绍该功能。该服务自身仅控制自动后台迁移。无论是否启用该服务，都将为带内请求同步迁移数据。

应仅针对测试目的或者在由于影子迁移而导致系统上的负载过大时禁用该服务。禁用时，将没有文件系统完成迁移。该服务的主要用途是允许调整专用于后台迁移的线程数。

影子迁移属性

表 11-23 影子迁移属性

属性	说明
Number of Threads（线程数）	专用于后台数据迁移的线程数。这些线程对于整个计算机是全局的，增加该数量可以增加并发性和迁移整体速度，但是会增加资源消耗（网络、I/O 和 CPU）。

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

SFTP 服务

SFTP（SSH File Transfer Protocol，SSH 文件传输协议）服务支持从 SFTP 客户机访问文件系统。不允许匿名登录，用户必须使用服务中配置的名称服务进行验证。

SFTP 属性

- Port (端口) (用于传入连接) – SFTP 侦听的端口。默认值是 218
- Permit root login (允许 root 登录) – 允许 root 用户进行 SFTP 登录。默认情况下处于关闭状态。
- Logging level (登录级别) – SFTP 日志消息的详细程度
- SFTP Keys (SFTP 密钥) – 用于 SFTP 验证的 RSA/DSA 公钥。文本注释可与密钥关联, 以帮助管理员跟踪添加原因。从 2011.1 软件发行版开始, 已经更改了 SFTP 的密钥管理以提高安全性。在创建 SFTP 密钥时, 需要包括具有有效用户分配的 "user" 属性。SFTP 密钥按用户分组, 并通过 SFTP 和用户的名称进行验证。建议重新创建不包括 user 属性的任何现有 SFTP 密钥, 不管这些密钥是否仍将进行验证。

有关更改服务属性的信息, 请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

SFTP 端口

SFTP 服务使用非标准端口号进行设备连接。这样可以避免与端口 22 的管理性 SSH 连接冲突。默认情况下, SFTP 端口为 218, 并且必须在连接之前在 SFTP 客户机上指定此端口。例如, 使用 SFTP 的 OpenSolaris 客户机将使用以下命令连接:

```
manta# sftp -o "Port 218" root@guppy
```

SFTP 日志

表 11-24 SFTP 日志

日志	说明
network-sftp:default	记录 SFTP 服务事件

配置 SFTP

▼ 允许对共享资源进行 SFTP 访问

1. 转到 "Configuration" (配置) -> "Services" (服务)
2. 检查 SFTP 服务是否已启用且处于联机状态。如果不是, 请启用该服务。

3. 在 "Shares" (共享资源) 屏幕中选择或添加一个共享资源。
4. 转到 "Protocols" (协议) 部分, 检查是否启用了 SFTP 访问。此处还可以设置访问模式 (读取/读取+写入)。

▼ 为远程访问配置 SFTP 服务

1. 创建一个具有适当管理员角色的本地用户或网络用户 (LDAP 或 NIS)。(请参见第 7 章 用户配置。)
2. 通过在 Solaris 主机/客户机上输入 `ssh-keygen -t dsa` 命令, 生成一个 SSH 验证密钥。
3. 输入要在其中存储密钥的文件的名称。
4. 如果需要, 请输入密码, 或将此字段保留空白, 直接登录 SFTP 共享资源。显示该位置以供密钥使用。密钥与以下内容类似:
5. `:ssh-dss AAAAB3NzaC1kc3MAAACBAPMMs5h8UWk1NPf/VJDDEo0OAwt+s6iZxkCmmrgAmLfTX9izWk+`
6. `:bsvNldOIXN/6EgkusLjo/+UaEt5+704vMHCIraq3AlVHLS5tVjeX3iCs+fDo0qwXZg3Brh8QBAaWk3`
7. `:ywr2osull1tHh4v/HwEAHZq5mVWXav0pO3bgmxl0/+VAAAFQDIJxnm52DfyEdQQMTY+jRVvzGwMQA`
8. `:AAIAhTP6Ey+2gGFICKkvUofsc04d8pbqH8duE9P6Y88s0+opuj52GkAdRUt2fRrdM9Cf3h4lIOc8Bw9`
9. `:bZlBzrCKBNWBudZG56tsfLdilW6vS6gxKrmL2v7fSp9WYPsxZGhOLfU29zW4n2WVcVHbGyFEoVe+taq`
10. `:aq+AYJaWoHnjZL1/LpQAAAIaOLc8+uc3hDOcK3pAkYdg8b2rYIGOAzu4py0rq24DGPeVHd5h5jbe4p`
11. `:WDM70uYqGCOPYiOKeEoMNJpczRX5qjl+BfoUY4sH24WWwsKkT8XX9PUAa0WT+7axEqg2N6YelaTJ95J`
12. `:vMaj6E7HkAlra2Sj2H/LSDktL42UL+j1Wx5A== username sunray`
13. 转到 "Configuration" (配置) > "Services" (服务) > "SFTP" 在密钥下方, 单击加号 (+)。

14. 在新密钥窗口中，选择 DSA。
15. 仅复制关键部分（上述示例中，以 AAAA 开头，以 Wx5A== 结尾），将其粘贴在密钥字段中。输入用户名称并添加注释作为提示信息。
16. 注意：该密钥不得包含空格。
17. 转到 "Shares"（共享资源） > "Shares"（共享资源），单击加号 (+) 创建文件系统。
18. 在 "Create Filesystem"（创建文件系统）窗口中，输入文件系统名称（例如 sftp），将共享资源权限更改为 "Read/ Write"（读取/写入），单击 "Apply"（应用）。
19. 单击铅笔图标以设置共享资源属性。（请参见[第 12 章 共享资源、项目和模式](#)。）
20. 要访问该共享资源，请使用下面这些示例中显示的 `sftp` 命令：
21. `: sftp -o "port=218" <username> 10.x.x.151:/export/sftp`
22. `: Connecting to 10.x.xx.151...`
23. `: Changing to: /export/sftp`
24. `: sftp>`
25. `: Example with -v option:`
26. `: sftp -v -o "IdentityFile=/home/<username>/.ssh/id_dsa" -o "port=218"`
27. `: root 10.x.xx.151:/export/sftp`

SRP 服务

在设备上配置 LUN 时，可通过一个 SCSI 远程协议 (SCSI Remote Protocol, SRP) 目标导出该数据。SRP 服务允许启动器使用 SRP 协议来访问目标。

有关 SRP 目标和启动器的详细信息，请参见[第 6 章 存储区域网络配置](#)。

有关管理 SRP 目标的示例，请参见[第 6 章 存储区域网络配置](#)。

TFTP 服务

普通文件传输协议 (Trivial File Transfer Protocol, TFTP) 是一种传输文件的简单协议。TFTP 设计小巧且很容易实施，因此缺乏常规 FTP 的大部分功能。TFTP 仅从远程

服务器读取/向其写入文件（或邮件）。它不能列出目录，目前没有针对用户身份验证的规定。

TFTP 属性

表 11-25 TFTP 属性

属性	说明
Default Root Directory（默认根目录）	TFTP 登录位置。默认值为 "/export"，指向共享资源分层结构的顶层。所有用户在成功进行 TFTP 服务验证后都将登录到此位置

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

配置 TFTP

▼ 允许对共享资源进行 TFTP 访问

1. 转到 "Configuration"（配置）-> "Services"（服务）
2. 检查 TFTP 服务是否已启用且处于联机状态。如果不是，请启用该服务。
3. 在 "Shares"（共享资源）屏幕中选择或添加一个共享资源。
4. 转到 "Protocols"（协议）部分，检查是否启用了 TFTP 访问。此处还可以设置访问模式（读取/读取+写入）。

病毒扫描服务

病毒扫描服务将在文件系统级别扫描病毒。通过任何协议访问文件时，病毒扫描服务都会先扫描文件；如果发现病毒，就会拒绝访问并隔离文件。使用最新的病毒定义扫描了文件之后，在下次修改文件之前不会再扫描该文件。具有缓存文件数据或者由 NFSv4 服务器委托了读取特权的 NFS 客户机访问的文件不能立即隔离。

病毒扫描属性

表 11-26 病毒扫描属性

属性	说明
Maximum file size to scan (要扫描的最大文件大小)	将不扫描大于此大小的文件以避免严重影响性能。这些大文件本身不可能是可执行文件（如数据库文件），因此不太可能对易受攻击的客户机带来风险。默认值为 1GB。
Allow access to files that exceed maximum file size (允许访问超过最大文件大小的文件)	默认情况下启用；此选项允许访问大于最大扫描大小的文件（因为太大而未在返回客户机前扫描的文件）。安全要求更严格的站点上的管理员可以选择禁用此选项并增加最大文件大小，以确保对所有可访问文件进行病毒扫描。

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

病毒扫描文件扩展名

本节介绍如何控制扫描的文件。默认值 "*" 将使所有文件被扫描。扫描所有文件可能会影响性能，所以可以指定扫描一部分文件。

例如，要扫描包括 zip 文件在内的所有高风险文件，但不扫描其名称符合模式 "data-archive*.zip" 的文件，可以配置以下设置：

表 11-27 病毒扫描文件扩展名

操作	模式
Scan (扫描)	exe
Scan (扫描)	com
Scan (扫描)	bat
Scan (扫描)	doc
Scan (扫描)	zip
Don't Scan (不扫描)	data-archive*.zip
Don't Scan (不扫描)	*

注意：您必须使用 "Don't Scan *" (不扫描 *) 排除未在扫描列表中显式包含的其他所有文件类型。不会扫描名为 "file.name.exe.bat.jpg123" 的文件，因为只有名称的 "jpg123" 部分即扩展名会与规则进行比较。

请不要在包含设置之前使用排除设置。例如，请不要在包含设置之前使用 "Don't Scan *"（不扫描 *）设置，因为那样将排除它之后的所有文件类型。以下示例不扫描任何文件：

表 11-28 病毒扫描操作

操作	模式
Don't Scan（不扫描）	*
Scan（扫描）	exe
Scan（扫描）	com
Scan（扫描）	bat
Scan（扫描）	doc
Scan（扫描）	zip
Don't Scan（不扫描）	data-archive*.zip

扫描引擎

在此部分中，指定要使用的扫描引擎。扫描引擎是一种外部第三方病毒扫描服务器，设备使用 ICAP（Internet Content Adaptation Protocol，Internet 内容修改协议，RFC 3507）与该服务器通信以进行文件扫描。

表 11-29 扫描引擎属性

属性	说明
Enable（启用）	使用此扫描引擎
Host（主机）	扫描引擎服务器的主机名或 IP 地址
Maximum Connections（最大连接数）	最大并发连接数。有些扫描引擎在连接数限制为 8 个时会运行得更好。
Port（端口）	用于扫描引擎的端口

病毒扫描日志

表 11-30 病毒扫描日志

日志	说明
vscan	病毒扫描服务的日志

配置病毒扫描

▼ 配置共享资源的病毒扫描

- 1. 转到 "Configuration" (配置) -> "Services" (服务) -> "Virus Scan" (病毒扫描) 。
- 2. 设置所需的属性。
- 3. 应用/提交配置。
- 4. 转到 "Shares" (共享资源) 。
- 5. 编辑文件系统或项目。
- 6. 选择 "General" (常用) 选项卡。
- 7. 启用 "Virus scan" (病毒扫描) 选项。

NIS 服务

网络信息服务 (Network Information Service, NIS) 是一种用于集中管理的名称服务。设备可以充当用户和组的 NIS 客户机，以便：

- NIS 用户可以登录[“FTP 服务” \[206\]](#)和[“HTTP 服务” \[208\]](#)。
- 可向 NIS 用户授予设备管理特权。设备使用自己的特权设置补充 NIS 信息。

请注意，0-99（含两端）的 UID 和 GID 由操作系统供应商预留用于将来的应用程序。分层产品的最终系统用户或供应商使用它们不受支持并且可能会导致将来的应用程序出现安全相关问题。

NIS 属性

表 11-31 NIS 属性

属性	说明
Domain (域)	要使用的 NIS 域
Server(s): Search using broadcast (服务器：使用广播搜索)	设备将发送 NIS 广播以查找该域的 NIS 服务器

属性	说明
Server(s): Use listed servers (服务器：使用列出的服务器)	NIS 服务器主机名或 IP 地址

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

设备将连接到使用广播列出或找到的第一个 NIS 服务器；如果该服务器停止响应，则切换到下一个。

NIS 日志

表 11-32 NIS 日志

日志	说明
network-nis-client:default	NIS 客户机服务日志
appliance-kit-nsswitch:default	设备名称服务的日志，通过它进行 NIS 查询
system-identity:domain	设备域名配置器的日志

配置 NIS

▼ 从 NIS 添加设备管理员

如果您在 NIS 中有现有用户，并且该用户想要使用其 NIS 凭证登录并管理设备：

1. 转到 "Configuration" (配置) -> "Services" (服务) -> "NIS"
2. 设置 NIS 域和服务器属性。
3. 应用/提交配置。
4. 转到 "Configuration" (配置) -> "Users" (用户)
5. 添加类型为 "directory" 的用户
6. 将用户名设置为其 NIS 用户名
7. 使用[第 7 章 用户配置](#)中的说明继续，为用户添加授权。

LDAP 服务

LDAP (Lightweight Directory Access Protocol, 轻量目录访问协议) 是一种目录服务, 用于集中管理用户、组、主机名和其他资源 (称为对象)。此服务在设备上充当 LDAP 客户端, 以便:

- LDAP 用户可以登录“FTP 服务” [206]和“HTTP 服务” [208]。
- 可使用 LDAP 用户名 (而不是数字 ID) 配置针对共享资源的根目录 ACL。
- 可向 LDAP 用户授予设备管理特权。设备使用自己的特权设置补充 LDAP 信息。
- LDAP 服务器证书可以自签名。
- 您无法提供可信 CA 证书列表; 每个证书必须分别由设备管理员接受。
- 当 LDAP 服务器证书过期时, 您必须从列表中将服务器删除, 并重新添加以接受其新的证书。

请注意, 0-99 (含两端) 的 UID 由操作系统供应商预留用于将来的应用程序。分层产品的最终系统用户或供应商使用它们不受支持并且可能会导致其他应用程序出现安全问题。

LDAP 属性

有关适合您环境的设置, 请咨询 LDAP 服务器管理员。

- Protect LDAP traffic with SSL/TLS (使用 SSL/TLS 保护 LDAP 通信) – 使用 TLS (Transport Layer Security, 传输层安全, SSL 的后代协议) 建立到 LDAP 服务器的安全连接
- Base search DN (基础搜索 DN) – 提供基础对象的可识别名称, 作为目录搜索的起点。
- Search scope (搜索范围) – 相对于基本对象, 定义在 LDAP 目录中搜索的对象。可将搜索结果限制为仅显示基本搜索对象正下方的对象 (one-level), 也可在搜索结果中包括基本搜索对象下方的任何对象 (subtree)。默认设置为 one-level。
- Authentication method (验证方法) – 用于向 LDAP 服务器验证设备的方法。设备支持 Simple (RFC 4513)、SASL/DIGEST-MD5 和 SASL/GSSAPI 验证。如果使用 Simple 验证方法, 则应启用 SSL/TLS, 使用户的 DN 和密码不以纯文本格式发送。在使用 SASL/GSSAPI 验证方法时, 仅可使用自绑定凭证级别。
- Bind credential level (绑定凭证级别) – 用于向 LDAP 服务器验证设备的凭证。
- * Anonymous (匿名) 为设备提供仅能访问可用于所有人的数据的权限。
- * Proxy (代理) 指示通过指定账户绑定的服务。
- * Proxy DN (代理 DN) – 用于代理验证的账户的标识名。
- * Proxy Password (代理密码) – 用于代理验证的账户的密码。
- * Self (自我) – 使用用户的身份与凭证来自我验证设备。自我验证仅能与 SASL/GSSAPI 验证方法一起使用。

- Schema definition（模式定义）– 设备使用的模式。管理员可使用此属性覆盖用户、组和网络组的默认搜索描述符、属性映射和对象类映射。有关详细信息，请参见“LDAP 服务” [225]。
- Servers（服务器）– 要使用的 LDAP 服务器列表。如果仅指定了一个服务器，设备将仅使用该服务器；如果该服务器发生故障，则 LDAP 服务将无法使用。如果指定多个服务器，则会随时使用任何正常运行的服务器，而没有任何优先级。如果任何服务器发生故障，则会使用列表中的其他服务器。LDAP 服务将持续可用，除非指定的所有服务器都发生故障。

LDAP 定制映射

要在 LDAP 目录中查找用户和组，设备应使用搜索描述符，且必须知道哪些对象类与用户和组相对应，哪些特性与所需属性相对应。默认情况下，设备使用 RFC 2307 指定的对象类（*posixAccount* 和 *posixGroup*）以及如下列表中所示的默认搜索描述符，但您可针对不同环境进行定制。下面示例中使用的基本搜索 DN 为 *dc=example,dc=com*：

表 11-33 LDAP 定制映射

搜索描述符	默认值	示例
Users（用户）	<i>ou=people,base search DN</i>	<i>ou=people,dc=example,dc=com</i>
groups（组）	<i>ou=group,base search DN</i>	<i>ou=group,dc=example,dc=com</i>
netgroups（网络组）	<i>ou=netgroup,base search DN</i>	<i>ou=netgroup,dc=example,dc=com</i>

可使用 "Schema definition"（模式定义）属性定制所使用的搜索描述符、对象类和特性。要覆盖默认搜索描述符，请输入您想要使用的完整 DN。设备将按原样使用此值，并忽略 "Base search DN"（基础搜索 DN）和 "Search scope"（搜索范围）属性的值。要覆盖用户、组和网络组的属性和对象，请选择相应的选项卡（"Users"（用户）、"Groups"（组）或 "Netgroups"（网络组）），然后使用 *default = new* 语法指定映射，其中 *default* 是默认值，*new* 是您要使用的值。例如：

- 要使用 *unixaccount* 来代替 *posixAccount* 作为用户对象类，请在 "Users"（用户）选项卡上的 "Object class mappings"（对象类映射）中输入 *posixAccount = unixaccount*。
- 要使用 *employeenumber* 来代替 *uid* 作为用户对象属性，请在 "Users"（用户）选项卡上的 "Attribute mappings"（属性映射）中输入 *uid = employeenumber*。
- 要使用 *unixgroup* 来代替 *posixGroup* 作为组对象类，请在 "Groups"（组）选项卡上的 "Object class mappings"（对象类映射）中输入 *posixGroup = unixgroup*。
- 要使用 *groupaccount* 来代替 *cn* 作为组对象属性，请在 "Groups"（组）选项卡上的 "Attribute mappings"（属性映射）中输入 *cn = groupaccount*。

以下列出了可能要映射的对象类和属性：

- 类 :
- * posixAccount
- * posixGroup
- * shadowAccount
- 属性 – 用户 :
- * uid
- * uidNumber
- * gidNumber
- * gecos
- * homeDirectory
- * loginShell
- * userPassword
- 属性 – 组 :
- * uid
- * memberUid
- * cn
- * userPassword
- * gidNumber
- * member
- * uniqueMember
- * memberOf
- * isMemberOf

LDAP 日志

以下为例日志。

表 11-34 LDAP 日志

日志	说明
appliance-kit-nsswitch:default	设备名称服务的日志，通过它进行 LDAP 查询

配置 LDAP

▼ 添加设备管理员

要使现有的 LDAP 用户日志用于 LDAP 身份凭证及管理设备，请执行以下过程：

1. 在 "Configuration" (配置) => "Services" (服务) => "LDAP" 页面中, 输入要使用的属性。有关可用属性的信息, 请参见[“LDAP 属性” \[225\]](#)。
2. 要应用选择的属性, 请单击 "Apply" (应用) 或单击 "Revert" (恢复) 重新开始。
3. 要添加 LDAP 服务器, 请在服务器部分单击添加  图标。有关服务器的信息, 请参见[“LDAP 属性” \[225\]](#)中的服务器部分。
4. 要配置 LDAP 服务器, 请在新 LDAP 服务器框中, 输入 LDAP 服务器地址, 并选择要使用的 LDAP 证书源。对于 "Certificate" (证书) 源, 选择 "Server" (服务器) 可搜索当前服务器, 检索证书 (以不安全的方式), 并用于未来验证后来提供的证书。
5. 在 "Configuration" (配置) => "Users" (用户) 页面上, 使用 LDAP 用户名根据需要添加用户。有关添加用户的信息, 请参见[第 7 章 用户配置](#)。

Active Directory

使用 Active Directory 服务可访问 Microsoft Active Directory 数据库, 其中存储有关用户、组、共享资源和其他共享对象的信息。该服务有两种模式: 域模式和工作组模式, 它们指示[“SMB” \[191\]](#)用户的验证方式。在域模式中运行时, [“SMB” \[191\]](#)客户机通过 AD 域控制器进行验证。在工作组模式中, [“SMB” \[191\]](#)客户机作为本地用户进行本地验证。有关本地用户的详细信息, 请参见[“用户”](#)。

Active Directory 属性

Active Directory 加入域

默认情况下, 如果帐户未存在于 Active Directory 中, 则在加入域的操作过程中, 会自动在计算机帐户的默认容器中 (cn=Computers) 为系统创建一个计算机信任帐户。允许以下用户执行域加入:

- 域管理员。可以将任意数量的系统加入域, 并可将计算机信任帐户放置在任何容器中。
- 具有对一个或多个组织单元的授权的受委托管理员。可以将任意数量的系统加入域, 并可以将计算机帐户位置指定在他们负责的组织单元中。
- 具有由管理员预设置的计算机帐户的普通用户。可以根据管理员的预授权将系统加入域。
- 普通用户。通常被授权加入有限数量的系统。

以下为加入 Active Directory 域的可用属性:

- Active Directory Domain (Active Directory 域) – Active Directory 域的全限定名称或 NetBIOS 名称
- User (用户) – 具有可在 Active Directory 中创建计算机帐户的凭证的 AD 用户
- Password (密码) – 管理用户的密码
- Additional DNS Search Path (其他 DNS 搜索路径) – 指定此可选属性后, 除了主 DNS 域和 Active Directory 域之外, 还会针对此域解析 DNS 查询
- Organizational Unit (组织单元) – 指定将在其中创建系统的计算机信任帐户的备选组织单元。组织单元将指定为包含一个或多个“名称 – 值”对的逗号分隔列表, 使用域相对标识名 (distinguished name, DN) 格式, 例如, ou=innerOU、ou=outerOU。
- Use Pre-created Account (使用预先创建的帐户) – 如果系统帐户已存在, 而且指定的组织单元不是帐户所在的单元, 请使用预先创建的帐户。

Active Directory 加入工作组

以下列表介绍了用于加入工作组的可配置属性。

- Windows Workgroup (Windows 工作组) – 一个工作组

有关更改服务属性的信息, 请参见“[使用 BUI 配置服务](#)” [179]和“[使用 CLI 配置服务](#)” [181]。CLI 属性名称是上面列出的名称的短版本。

Active Directory 域和工作组

通过加入域或工作组来修改服务, 而不是直接启用和禁用服务。加入域包括在给定的 Active Directory 域中为设备创建帐户。帐户名称最多为 15 个字符, 在注册到 Active Directory 域的所有其他名称中唯一。否则, 可能会出现相同命名设备冲突, 导致功能出现问题。创建计算机帐户后, 设备便可在数据库中安全地查询有关用户、组和共享资源的信息。

加入工作组会隐式离开 Active Directory 域, 并且存储在 Active Directory 数据库中的“SMB” [191]客户机将无法连接到共享资源。

如果将 Kerberos 领域配置为支持基于 Kerberos 的 NFS, 则无法将系统配置为加入 Active Directory 域。

Active Directory LDAP 签名

没有用于 LDAP 签名的配置选项, 因为在与域控制器进行通信时会自动协商该选项。存储设备与域控制器之间的通信使用 LDAP 签名, 而 SMB 客户机与存储设备之间的通信使用 SMB 签名。

Active Directory Windows Server 2012 支持

Windows Server 2012 在软件版本 2011.1.5 及更高版本中完全受支持。

Active Directory Windows Server 2008 支持

表 11-35 Active Directory Windows Server 2008 支持

Windows 版本	支持的软件版本	解决方法
Windows Server 2003	全部	无
Windows Server 2008 SP1	2009.Q2 3.1 和更低版本	根据需要应用 KB957441 中所述的热修复程序，请参见 B 部分。
	2009.Q2 4.0 - 2011.1.1	必须应用 KB951191 中所述的热修复程序，并根据需要应用 KB957441 中所述的热修复程序，请参见 A 和 B 部分。
	2011.1.2 和更高版本	必须应用 KB951191 中所述的热修复程序，请参见 A 部分。
Windows Server 2008 SP2	2009.Q2 4.0 - 2011.1.1	请参见 C 部分。
	2011.1.2 和更高版本	无
Windows Server 2008 R2	2009.Q2 4.0 - 2011.1.1	请参见 C 部分。
	2011.1.2 和更高版本	无

Active Directory Windows Server 2008 支持部分 A : Kerberos 发布 (KB951191)

- 如果您升级至 2009.Q2.4.0 或更高版本且您的 Windows 2008 域控制器正在运行 Windows Server 2008 SP2 或 R2，则无需执行任何操作。
- 如果您升级至 2009.Q2.4.0 或更高版本且您的 Windows 2008 域控制器正在运行 Windows Server 2008 SP1，则必须应用 KB951191 中所述的热修复程序或安装 Windows 2008 SP2。

Active Directory Windows Server 2008 支持部分 B : NTLMv2 发布 (KB957441)

- 以下内容仅在设备运行的软件版本低于 2011.1.2 时才适用：
- 如果您的域控制器正在运行 Windows Server 2008 SP1，也应该应用 <http://support.microsoft.com/kb/957441/> (<http://support.microsoft.com/kb/957441/>) 中所

述的热修复程序，它可解决会阻止设备使用其默认 LMCompatibilityLevel 设置加入域的 NTLMv2 问题。

- 如果 Windows 2008 SP1 域控制器上的 LMCompatibilityLevel 设置为 5，则必须安装此热修复程序。在应用该热修复程序之后，您必须按 KB957441 中所述创建和设置新的注册表项。
- 如果要升级到 2011.1.2 或更高版本，则不需要上面所述的热修复程序。

Active Directory Windows Server 2008 支持部分 C：NTLMv2 的注释

- 以下内容仅在设备运行的软件版本低于 2011.1.2 时才适用：如果您的域控制器正在运行 Windows Server 2008 SP2 或 R2，则不需要应用该热修复程序，但必须按 KB957441 中所述应用注册表设置。
- 如果要升级到 2011.1.2 或更高版本，则不需要执行任何操作。

使用 BUI 配置 Active Directory

▼ 加入域

1. 在“SMB” [191]上下文中配置 Active Directory 站点。（可选）
2. 在“SMB” [191]上下文中配置首选域控制器。（可选）
3. 启用“NTP” [243]或确保设备与域控制器的时钟同步（相差在五分钟之内）。
4. 确保您的“DNS” [239]基础结构已正确委托给 Active Directory 域，或将您域控制器的 IP 地址作为附加的名称服务器添加到“DNS” [239]上下文中。
5. 配置 Active Directory 域、管理用户和管理密码。
6. 应用/提交配置。

▼ 加入工作组

1. 配置工作组名称。
2. 应用/提交配置。

使用 CLI 配置 Active Directory

为了演示 CLI 界面，以下示例将查看现有配置，加入工作组，加入域。

▼ 示例 – 使用 CLI 配置 Active Directory

1. 查看现有配置。

```
twofish:> configuration services ad
twofish:configuration services ad> show
Properties:
    <status> = online
    mode = domain
    domain = eng.fishworks.com

Children:
    domain => Join an Active Directory domain
    workgroup => Join a Windows workgroup
```

2. 请注意，设备当前在 "eng.fishworks.com" 域中运行。以下是离开该域并加入一个工作组的示例。

```
twofish:configuration services ad> workgroup
twofish:configuration services ad workgroup> set workgroup=WORKGROUP
twofish:configuration services ad workgroup> commit
twofish:configuration services ad workgroup> done
twofish:configuration services ad> show
Properties:
    <status> = disabled
    mode = workgroup
    workgroup = WORKGROUP
```

3. 以下是配置站点和首选域控制器以便加入另一域的示例。

```
twofish:configuration services ad> done
twofish:> configuration services smb
twofish:configuration services smb> set ads_site=sf
twofish:configuration services smb> set pdc=192.168.3.21
twofish:configuration services smb> commit
twofish:configuration services smb> show
Properties:
    <status> = online
    lmauth_level = 4
    pdc = 192.168.3.21
    ads_site = sf
twofish:configuration services smb> done
```


4. 以下是在配置属性之后加入新域的示例。加入 AD 域后，每次提交节点时必须设置用户和密码。

```
twofish:> configuration services ad
twofish:configuration services ad> domain
twofish:configuration services ad domain> set domain=fishworks.com
twofish:configuration services ad domain> set user=Administrator
twofish:configuration services ad domain> set password=*****
twofish:configuration services ad domain> set searchdomain=it.fishworks.com
twofish:configuration services ad domain> commit
twofish:configuration services ad domain> done
twofish:configuration services ad> show
Properties:
    <status> = online
    mode = domain
    domain = fishworks.com
```

身份映射服务

身份映射服务使用传统 Unix UID（和 GID）和 Windows SID 同时管理 Windows 和 Unix 用户身份。有关使用 BUI 和 CLI 进行身份映射的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。

身份映射属性

身份映射服务可创建 SID、UID 和 GID 之间的映射数据库并对其进行维护。有三种不同的映射方法可用，如果映射可用于指定的身份，服务将会创建临时映射。以下映射模式可用：

基于身份映射规则的映射

基于规则的映射方法包括创建各种按名称映射身份的规则。这些规则在 Windows 身份与 Unix 身份之间建立对等关系。

基于身份映射目录的映射

基于目录的映射涉及为[“LDAP” \[225\]](#)或[“Active Directory” \[228\]](#)对象加注有关如何将身份映射到对应平台上的对等身份的信息。在使用基于目录的映射时，必须指定以下属性：

- AD 属性 – Unix User Name（Unix 用户名称） – AD 数据库中对等 Unix 用户的名称

- AD 属性 – Unix Group Name (Unix 组名称) – AD 数据库中对等 Unix 组的名称
- 本地 LDAP 属性 – Windows User Name (Windows 用户名称) – LDAP 数据库中对等 Windows 身份的名称

CLI 属性名称是上面列出的名称的短版本。

有关扩充 [“Active Directory” \[228\]](#) 或 [“LDAP” \[225\]](#) 架构的信息，请参见《Solaris CIFS Administration Guide》(《Solaris CIFS 管理指南》) 中的 “Managing Directory-Based Identity Mapping for Users and Groups (Task Map)” (“管理用户和组的基于目录的身份映射 (任务列表)”) 一节。

身份映射 IDMU

Microsoft 提供了一项称为 “Identity Management for Unix (IDMU)” 的功能。Windows Server 2003 中提供了此软件，并且 Windows Server 2003 R2 和更高版本捆绑了此软件。此功能是之前称为 “Services for UNIX” 的功能的一部分 (采用非捆绑形式)。

IDMU 的主要用途是支持将 Windows 作为 NIS/NFS 服务器。IDMU 在 “Active Directory Users and Computers” (Active Directory 用户和计算机) 用户界面上添加了 “UNIX Attributes” (UNIX 属性) 面板，管理员可在其中指定一些与 UNIX 相关的参数：UID、GID、登录 shell、主目录以及组的类似参数。这些参数在 AD 中通过与 RFC2307 相似 (但不完全相同) 的模式以及 NIS 服务设置。

选择 IDMU 映射模式时，身份映射服务会使用这些 Unix 属性，以在 Windows 身份与 Unix 身份之间创建映射。此方法与基于目录的映射非常相似，只是身份映射服务会查询 IDMU 软件创建的属性模式，而不是允许使用定制模式。使用此方法时，无法使用任何其他基于目录的映射。

身份映射规则

本页介绍使用以下属性创建映射：

- Mapping Type (映射类型) – 允许或拒绝凭证。有关更多信息，请参阅[“拒绝映射” \[233\]](#)。
- Mapping Direction (映射方向) – 映射方向。一个映射可以向两个方向映射凭证：仅从 Windows 到 Unix 或仅从 Unix 到 Windows 有关更多信息，请参阅[“映射规则方向符号” \[233\]](#)。
- Windows Domain (Windows 域) – Windows 身份的 Active Directory 域。
- Windows Identity (Windows 身份) – Windows 身份的名称。
- Unix Identity (Unix 身份) – Unix 身份的名称。
- Unix Identity Type (Unix 身份类型) – Unix 身份的类型，用户或组。

拒绝映射

拒绝映射规则可阻止用户从身份映射服务获取任何映射，包括临时 ID。您可以为 Windows 用户和 Unix 用户创建全域性或用户特定拒绝映射。例如，您可以创建一个映射，以禁止组 "guest" 中的所有 Unix 用户访问“SMB” [191]共享资源。您无法创建与其他映射冲突的拒绝映射。

映射规则方向符号

创建基于名称的映射之后，以下符号可指示每个规则的语义。

- align="center"| – 将 Windows 身份映射到 Unix 身份，并将 Unix 身份映射到 Windows 身份
- align="center"| – 将 Windows 身份映射到 Unix 身份
- align="center"| – 将 Unix 身份映射到 Windows 身份
- align="center"| – 阻止 Windows 身份获取凭证
- align="center"| – 阻止 Unix 身份获取凭证

如果图表是灰色而非黑色（、、、、），该规则将匹配无法解析的 Unix 身份。

身份映射映射

"Mappings" (映射) 页面显示各个身份如何使用当前的一组规则进行映射。通过指定 Windows 实体或 Unix 实体，该实体将映射到其在对应平台上的相应身份。"User Properties" (用户属性) 和 "Group Properties" (组属性) 部分的结果信息显示有关映射身份的信息，包括映射源。本页介绍使用 "Show" (显示) 和 "Flush" (刷新) 按钮查看和删除现有映射。

身份映射日志

本页显示最近活动的日志。

身份映射最佳实践

- 仅当您想让同一用户同时以“NFS” [185]和“SMB” [191]客户机访问一组通用文件时，才适合配置细化身份映射规则。如果“NFS” [185]和“SMB” [191]客户机要访问不相交的文件系统，则不需要配置任何身份映射规则。

- 重新配置身份映射服务对活动的“SMB” [191]会话没有任何影响。已连接的用户将保持连接，且其先前的名称映射可用于授权访问其他共享资源（最多 10 分钟）。要防止未经授权的访问，您必须在导出共享资源之前配置映射。
- 身份映射的安全性取决于其与您的目录服务的同步程度。例如，如果您创建了一个基于名称的映射来拒绝访问某特定用户，而该用户的名称发生了更改，则该映射不会再拒绝访问该用户。
- 您只能为每个 Windows 域设置一个双向映射，来将该 Windows 域中的所有用户映射到所有 Unix 身份。如果您想创建多个全域性规则，请确保指定这些规则仅从 Windows 映射到 Unix。
- 尽可能使用 IDMU 映射模式，而非基于目录的映射。

身份映射概念

“SMB” [191]服务使用身份映射服务来关联 Windows 和 Unix 身份。在“SMB” [191]服务对用户进行验证时，它会使用身份映射服务将用户的 Windows 身份映射到相应的 Unix 身份。如果某 Windows 用户没有 Unix 身份，该服务会使用临时的 UID 和 GID 生成临时身份。这些映射可允许“SMB” [191]和“NFS” [185]客户机同时导出和访问一个共享资源。将 Windows 身份与 Unix 身份相关联之后，“NFS” [185]和“SMB” [191]客户机便可共享相同身份，从而可访问相同的一组文件。

在 Windows 操作系统中，访问令牌包含登录会话的安全信息，并标识用户、用户组以及用户特权。管理员在工作组或 SAM 数据库（在“Active Directory” [228]域控制器中进行管理）中定义 Windows 用户和组。每个用户和组都有一个 SID。SID 可唯一标识一个主机和一个本地域以及所有可能的 Windows 域中的用户或组。

Unix 根据用户验证和文件权限创建用户凭证。管理员在本地密码和组文件或者名称或目录服务（例如“NIS” [223]和“LDAP” [225]）中定义 Unix 用户和组。每个 Unix 用户和组都有一个 UID 和一个 GID。通常，UID 或 GID 可唯一标识单个 Unix 域中的用户或组。但是，这些值在域之间并不是唯一的。

身份映射区分大小写

Windows 名称不区分大小写，Unix 名称区分大小写。用户名 JSMITH、JSmith 和 jsmith 在 Windows 中是相同的名称，但在 Unix 中则是三个不同的名称。大小写对名称映射有不同影响，具体取决于映射方向。

- 为使从 Windows 到 Unix 的映射生成匹配项，Windows 用户名的大小写必须与 Unix 用户名的大小写匹配。例如，只有 Windows 用户名 "jsmith" 与 Unix 用户名 "jsmith" 匹配。Windows 用户名 "Jsmith" 则不匹配。
- 有一个例外是，当映射使用通配符 "*" 映射多个用户名时，从 Windows 到 Unix 的映射不要求大小写匹配。如果身份映射服务遇到将 Windows 用户 *@some.domain 映射到 Unix 用户 "*" 的映射，它会先搜索与 Windows 名称完全相同的 Unix 名称。如果找不到匹配项，该服务会将整个 Windows 名称转换为小写形式，并再次搜索匹配的 Unix 名称。例如，Windows 用户名 "JSmith@some.domain" 映射到

Unix 用户名 "jsmith"。如果在将 Windows 用户名转换为小写形式之后，该服务还是找不到匹配项，则该用户没有映射成功。您可以创建一个规则来匹配只有大小写不同的字符串。例如，您可以创建一个特定于用户的映射，以将 Windows 用户 "JSmith@sun.com" 映射到 Unix 用户 "jSmith"。否则，该服务会将临时 ID 分配给 Windows 用户。

- 为使从 Unix 到 Windows 的映射生成匹配项，大小写不必匹配。例如，Unix 用户名 "jsmith" 与含有 "JSMITH" 这几个字母（不考虑大小写）的任何 Windows 用户名均匹配。

映射持久性

在身份映射服务提供名称映射时，它会存储映射 10 分钟，之后映射便会过期。在 10 分钟的生命周期内，映射会在身份映射服务重新启动后保留。如果“SMB” [191]服务器在映射已过期后请求对用户进行映射，该服务会重新计算映射。

在 10 分钟的映射生命周期内，对映射或名称服务目录的更改不会影响现有连接。仅当客户机尝试连接共享资源且没有未过期的映射时，该服务才会计算映射。

身份映射全域性规则

全域性映射规则会将 Windows 域中的部分或全部名称与 Unix 名称相匹配。两个域的用户名必须完全匹配（大小写冲突除外，具体取决于前面讨论的规则）。例如，您可以创建一个双向规则，以将 "myDomain.com" 中的所有 Windows 用户与同名的 Unix 用户相匹配（反之亦然）。再例如，您可以创建一个规则，以将 "myDomain.com" 中组 "Engineering" 中的所有 Windows 用户映射到同名的 Unix 用户。您无法创建与其他映射冲突的全域性映射。

临时映射

如果没有适用于特定用户的基于名称的映射规则，将通过临时映射为该用户提供临时凭证，除非他们已被拒绝映射阻止。当拥有临时 Unix 名称的 Windows 用户在系统上创建一个文件时，使用“SMB” [191]访问该文件的 Windows 客户机将认为该文件归该 Windows 身份所有。但是，“NFS” [185]客户机将认为该文件归 "nobody" 所有。

身份映射示例

以下是在 CLI 中添加两个基于名称的规则示例。第一个示例在 Windows 用户与 Unix 用户之间创建了一个基于名称的双向映射。

```
twofish:> configuration services idmap
twofish:configuration services idmap> create
twofish:configuration services idmap (uncommitted)> set
```

```

windomain=eng.fishworks.com
twofish:configuration services idmap (uncommitted)> set winname=Bill
twofish:configuration services idmap (uncommitted)> set direction=bi
twofish:configuration services idmap (uncommitted)> set unixname=wdp
twofish:configuration services idmap (uncommitted)> set unixtype=user
twofish:configuration services idmap (uncommitted)> commit
twofish:configuration services idmap> list
MAPPING      WINDOWS ENTITY      DIRECTION    UNIX ENTITY
idmap-000     Bill@eng.fishworks.com      (U) ==       wdp (U)

```

下一个示例创建了一个拒绝映射，来阻止某域中的所有 Windows 用户获取凭证。

```

twofish:configuration services idmap> create
twofish:configuration services idmap (uncommitted)> list
Properties:
    windomain = (unset)
    winname = (unset)
    direction = (unset)
    unixname = (unset)
    unixtype = (unset)

twofish:configuration services idmap (uncommitted)> set
windomain=guest.fishworks.com
twofish:configuration services idmap (uncommitted)> set winname=*
twofish:configuration services idmap (uncommitted)> set direction=win2unix
twofish:configuration services idmap (uncommitted)> set unixname=
twofish:configuration services idmap (uncommitted)> set unixtype=user
twofish:configuration services idmap (uncommitted)> commit
twofish:configuration services idmap> list
MAPPING      WINDOWS ENTITY      DIRECTION    UNIX ENTITY
idmap-000     Bill@eng.fishworks.com      (U) ==       wdp (U)
idmap-001     *@guest.fishworks.com      (U) =>       "" (U)

```

配置身份映射

▼ 配置身份映射

1. 确保您至少加入一个活动目录域。有关活动目录的信息，请参见[“Active Directory” \[228\]](#)部分。
2. 在 "Configuration" (配置) => "Services" (服务) => "Identity Mapping" (身份映射) => "Properties" (属性) 页面上，选择要使用的映射模式。有关映射模式的信息，请参见[“属性” \[233\]](#)。
3. 如果选择基于目录的映射，必须配置其他属性。有关这些属性的更多信息，请参见[“基于目录的映射” \[233\]](#)。
4. 要保存设置，请单击 "Apply" (应用) 或单击 "Revert" (恢复) 重新开始。

- 5. 要创建映射，请单击 "Rules"（规则）。
- 6. 在 "Rules"（规则）页面上，请单击添加  图标。
- 7. 在 "Add Mapping Rule"（添加映射规则）框中，请输入所需的信息。有关更多信息，请参见[“规则” \[233\]](#)
- 8. 要保存设置，请单击 "Add"（添加）或单击 "Cancel"（取消）。创建映射后，它将显示在 "Rules"（规则）列表中。

▼ 查看或刷新映射

- 1. 要查看现有映射，在 "Configuration"（配置）=> "Services"（服务）=> "Identity Mapping"（身份映射）=> "Mappings"（映射）页面上，请输入所需的信息。有关映射的信息，请参见[“映射” \[233\]](#)。
- 2. 单击 "Show"（显示）。您指定的映射将会显示。
- 3. 要删除映射，请单击 "Flush"（刷新）。该映射将被删除。

DNS 服务

DNS（Domain Name Service，域名服务）客户机可以将 IP 地址解析为主机名（反之亦然），且在设备上始终处于启用状态。或者，对于无法使用 DNS 解析的主机名和地址，可以请求通过 NIS 和/或 LDAP（如果已配置并启用）解析辅助主机名。主机名解析在所有用户界面中使用，包括在《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的[“日志”](#)日志中（用于指示用户执行可审计操作的位置），以及在《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的[“Analytics”](#)中（用于以客户机为单位提供统计信息）。

DNS 客户机的可配置属性包括基本域名以及由 IP 地址指定的一系列服务器。您必须提供一个域名以及至少一个服务器地址；该服务器必须能够返回您所指定的域的 NS（NameServer，名称服务器）记录，但其自身不需要是该域的权威服务器。

DNS 属性

表 11-36 DNS 属性	
属性	说明
DNS Domain（DNS 域）	在执行部分主机名查找时最先搜索的域名。

属性	说明
DNS Server(s) (DNS 服务器)	一个或多个 DNS 服务器。必须使用 IP 地址。
Allow IPv4 non-DNS resolution (允许 IPv4 非 DNS 解析)	可以使用 NIS 和/或 LDAP (如果已配置并启用) 将 IPv4 地址解析为主机名以及将主机名解析为 IPv4 地址。
Allow IPv6 non-DNS resolution (允许 IPv6 非 DNS 解析)	可以使用 NIS 和/或 LDAP (如果已配置并启用) 将 IPv4 和 IPv6 地址解析为主机名以及将主机名解析为 IPv4 和 IPv6 地址。

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

配置 DNS

CLI 包括nslookup和getent hosts的内置命令，可用于测试主机名解析是否正常：

```
caji:> nslookup deimos
192.168.1.109  deimos.sf.fishworks.com
caji:> getent hosts deimos
192.168.1.109  deimos.sf.fishworks.com
```

DNS 日志

表 11-37 DNS 日志

日志	说明
network-dns-client:default	记录 DNS 服务事件

Active Directory 和 DNS

如果您打算使用[“Active Directory” \[228\]](#)，服务器必须能够解析域名称空间的 Active Directory 部分中的主机名和服务器记录。例如，如果您的设备位于域 example.com 中，且名称空间的 Active Directory 部分为 redmond.example.com，则您的名称服务器必须能够连接 example.com 的权威服务器，且必须将域 redmond.example.com 委托给为该域服务的一个或多个 Active Directory 服务器。这些是 Active Directory 的要求，而不是设备自身的要求。如果不满足这些要求，您将无法加入 Active Directory 域。

非 DNS 解析

DNS 是用于在主机名和 IP 地址之间进行映射的一种标准、高度可伸缩且可靠的企业级机制。使用正常工作的 DNS 服务器是最佳做法，通常可获得最佳结果。在某些环境中，可能有一部分主机只能在 NIS 或 LDAP 映射中进行解析。如果在您的环境中出现这种情况，请启用非 DNS 主机解析并配置适当的目录服务。如果使用 LDAP 进行主机解析，主机映射必须位于您数据库中的标准 DN 处：ou=Hosts,(Base DN)，且必须使用标准模式。如果网络组将该模式与 NFS 共享结合使用，则客户机系统采用的主机名解析机制可能需要与设备上配置的相同，否则 NFS 共享例外可能无法正常处理。

启用非 DNS 主机解析时，仍将使用 DNS。仅当无法使用 DNS 解析某地址或主机名时，才会使用 NIS（如果已启用），然后使用 LDAP（如果已启用）来解析该名称或地址。这样会产生混淆且似乎不一致的结果。您可以使用上述 `getent` CLI 命令验证主机解析结果。

强烈建议您不要使用这些选项。

无 DNS 运行

设备不支持无 DNS 运行，无 DNS 运行会造成不良的结果。多项功能在没有 DNS 的情况下无法正确运行，包括但不限于：

- 《Oracle ZFS Storage Appliance Analytics 指南》中的“Analytics”无法将客户机地址解析为主机名。
- “Active Directory” [228]功能无法正常运行（无法加入域）。
- 受 SSL 保护的“LDAP” [225]无法正常用于包含主机名的证书。
- 涉及发送电子邮件的警报和阈值操作只能发送到连接的子网上的邮件服务器，且所有地址都必须使用邮件服务器的 IP 地址进行指定。
- 一些操作可能会因主机名解析超时而比正常情况下花费的时间要长。

动态路由服务

RIP 和 RIPng 动态路由协议

RIP（Routing Information Protocol，路由信息协议）是距离向量动态路由协议，设备使用该协议基于从其他支持 RIP 的链路主机（通常为路由器）接收的消息自动配置最佳路由。设备针对 IPv4 支持 RIPv1 和 RIPv2，针对 IPv6 支持 RIPng。通过这些协议配置的路由在路由表中标记为“dynamic”（动态）类型。RIP 和 RIPng 分别侦听 UDP 端口 520 和 521。

动态路由日志

表 11-38 动态路由

日志	说明
network-routing-route:default	记录 RIP 服务事件
network-routing-ripng:quagga	记录 RIPng 服务事件

IPMP 服务

IPMP (Internet Protocol Network Multipathing, Internet 协议网络多路径) 允许将多个网络接口作为一个接口分组在一起, 从而提高网络带宽和可靠性 (接口冗余)。可在此部分配置某些属性。有关在 IPMP 组中配置网络接口的信息, 请参见[第 4 章 网络配置](#)。

IPMP 属性

表 11-39 IPMP 属性

属性	说明
Failure detection latency (故障检测延迟)	IPMP 声明某网络接口已发生故障并故障转移到其 IP 地址的时间
Enable fail-back (启用故障恢复)	允许服务与已修复的接口恢复连接

有关更改服务属性的信息, 请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

IPMP 日志

表 11-40 IPMP 日志

日志	说明
network-initial:default	记录网络配置过程

NTP 服务

可使用网络时间协议 (Network Time Protocol, NTP) 服务保持设备时钟准确无误。这对于在文件系统中记录准确的时间戳以及对于协议验证都非常重要。设备使用 UTC 时区记录时间。显示在 BUI 中的时间使用您的浏览器的时区偏移。

NTP 属性

表 11-41 NTP 属性

属性	说明	示例
multicast address (多播地址)	在此输入多播地址以自动查找 NTP 服务器	224.0.1.1
NTP server(s) (NTP 服务器)	输入一个或多个 NTP 服务器 (及其对应的验证密钥 (如果有)) , 以便设备直接联系	0.pool.ntp.org
NTP Authentication Keys (NTP 验证密钥)	输入一个或多个 NTP 验证密钥, 以便设备在验证 NTP 服务器的有效性时使用。请参见下面的“身份验证”一节。	Auth key: 10, Type: ASCII, Private Key: SUN7000 (验证密钥: 10, 类型: ASCII, 私钥: SUN7000)

有关更改服务属性的信息, 请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

NTP 验证

如果输入无效的配置, 将显示一条警告消息, 且不会提交配置。在以下情况下会发生这种情况:

- 使用了多播地址, 但找不到任何 NTP 响应。
- 使用了 NTP 服务器地址, 但该服务器无法正确响应 NTP。

NTP 验证

为防止恶意服务器的 NTP 网络欺骗攻击, NTP 采用私钥加密方案, 其中 NTP 服务器与客户机用于验证其身份的私钥相关联。这些私钥并不用于加密通信, 也不用于验证客户机 - 它们仅供 NTP 客户机 (即设备) 用于验证 NTP 服务器。要将私钥与 NTP 服务器相关联, 必须先指定该私钥。每个私钥都有关联的唯一整数以及类型和密钥。此类型可以是以下类型之一:

表 11-42 NTP 私钥和整数

类型	说明	示例
DES	DES 格式的 64 位十六进制数字	0101010101010101
NTP	NTP 格式的 64 位十六进制数字	8080808080808080
ASCII	包含 1 到 8 个字符的 ASCII 字符串	topsecret
MD5	包含 1 到 8 个字符的 ASCII 字符串（使用 MD5 验证方案）。	md5secret

指定密钥之后，即可将 NTP 服务器与特定的私钥相关联。对于给定的密钥，客户机与服务器的所有密钥编号、密钥类型和私钥值都必须匹配，NTP 服务器才能通过验证。

NTP BUI 时钟

BUI 屏幕右侧是设备的时间（服务器时间）和浏览器的时间（客户机时间）。如果 NTP 服务未处于联机状态，可以单击 "SYNC"（同步）按钮，将设备时间设置为与您的客户机浏览器时间匹配。

NTP 提示

如果您使用 SMB 共享文件系统，必须将客户机时钟与设备时钟同步（相差在五分钟之内），以免发生用户验证错误。确保时钟同步的一种方法是将设备和 SMB 客户机配置为使用相同的 NTP 服务器。

表 11-43 NTP 时钟同步

日志	说明
network-ntp:default	NTP 服务日志

使用 BUI 配置 NTP

要在 BUI 中添加 NTP 验证密钥，请单击加号图标并指定新密钥的密钥编号、类型和私钥值。添加密钥之后，它将在指定的每个 NTP 服务器旁边显示为一个选项。

▼ BUI 时钟同步

这会将设备时间设置为与您浏览器的时间匹配。

1. 禁用 NTP 服务。
2. 单击 "SYNC" (同步) 按钮。

使用 CLI 配置 NTP

在 configuration services ntp 下，使用 authkey 命令编辑授权：

```
clownfish:configuration services ntp> authkey
clownfish:configuration services ntp authkey>
```

在此上下文中，可使用 create 命令添加新密钥：

```
clownfish:configuration services ntp authkey> create
clownfish:configuration services ntp authkey-000 (uncommitted)> get
    keyno = (unset)
    type = (unset)
    key = (unset)
clownfish:configuration services ntp authkey-000 (uncommitted)> set keyno=1
    keyno = 1 (uncommitted)
clownfish:configuration services ntp authkey-000 (uncommitted)> set type=A
    type = A (uncommitted)
clownfish:configuration services ntp authkey-000 (uncommitted)> set key=coconuts
    key = ***** (uncommitted)
clownfish:configuration services ntp authkey-000 (uncommitted)> commit
clownfish:configuration services ntp authkey>
```

要通过 CLI 将验证密钥与服务器相关联，应将 serverkeys 属性设置为一系列值，其中的每个值是要与 servers 属性中对应的服务器相关联的一个密钥。如果服务器不使用验证，应对应的服务器密钥设置为 0。例如，要使用上面创建的密钥验证服务器 "gefilte" 和 "carp"：

```
clownfish:configuration services ntp> set servers=gefilte,carp
    servers = gefilte,carp (uncommitted)
clownfish:configuration services ntp> set serverkeys=1,1
    serverkeys = 1,1 (uncommitted)
clownfish:configuration services ntp> commit
clownfish:configuration services ntp>
```

要用密钥 1 验证服务器 "gefilte"、用密钥 2 验证服务器 "carp"，并用密钥 3 验证服务器 "dory"：

```
clownfish:configuration services ntp> set servers=gefilte,carp,dory
    servers = gefilte,carp,dory (uncommitted)
```

```
clownfish:configuration services ntp> set serverkeys=1,2,3
serverkeys = 1,2,3 (uncommitted)
clownfish:configuration services ntp> commit
clownfish:configuration services ntp>
```

要用密钥 1 验证服务器 "gefilte" 和 "carp", 并另外设置一个未经验证的 NTP 服务器 "dory", 请执行以下操作 :

```
clownfish:configuration services ntp> set servers=gefilte,carp,dory
servers = gefilte,carp,dory (uncommitted)
clownfish:configuration services ntp> set serverkeys=1,1,0
serverkeys = 1,1,0 (uncommitted)
clownfish:configuration services ntp> commit
clownfish:configuration services ntp>
```

回拨服务

"Phone Home" (回拨) 服务屏幕用于管理设备注册以及回拨远程支持服务。

- 注册操作将设备与[Oracle 自动服务请求 \(Oracle Auto Service Request, ASR\)](http://oracle.com/asr) (<http://oracle.com/asr>) 功能连接。Oracle ASR 为您的设备报告的特定问题自动建立服务请求 (Service Request, SR)。注册操作还会将您的设备与 My Oracle Support (MOS) 连接以检测更新通知。
- 回拨服务与 Oracle 支持系统进行通信来提供以下功能：
- 故障报告 – 系统向 Oracle 报告活动问题以获得自动服务响应。根据故障的性质，可以创建支持案例。有关这些事件的详细信息，请在《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“问题”中查看。
- 心跳 – 将向 Oracle 发送每日心跳消息来指示系统是否启动并且正在运行。当一个已激活系统长时间未发送心跳时，Oracle 支持系统可能会通知帐户的技术联系人。
- 系统配置 – 将向 Oracle 发送定期消息，描述当前软件和硬件版本和配置以及存储配置。这些消息中不传输任何用户数据或元数据。
- 支持包 – 必须启用回拨服务，才能将支持包上传至 Oracle 支持。有关详细信息，请参见《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“系统”。
- 更新通知 – 当 My Oracle Support (MOS) 上有新的可用软件更新时，创建警报。有关详细信息，请参见《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“软件更新通知”。

您必须注册才能使用回拨服务。

Oracle 单点登录帐户

您需要有效 Oracle 单点登录帐户用户名和密码来使用回拨服务的故障报告和心跳功能。转到 <http://support.oracle.com> (<http://support.oracle.com>) 并单击 "Register" (注册) 创建帐户。

回拨属性

有关更改服务属性的信息，请参见“[使用 BUI 配置服务](#)” [179]和“[使用 CLI 配置服务](#)” [181]。回拨服务在 CLI 中称为 `scrk`。

回拨 Web 代理

如果设备没有直接连接到 Internet，则可能需要配置 HTTP 代理，回拨服务可以通过该代理与 Oracle 通信。这些代理设置还将用于上载支持包。有关支持包的详细信息，请参见《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“系统”。

表 11-44 回拨 Web 代理设置

属性	说明
Use proxy (使用代理)	通过 web 代理连接
Host/port (主机/端口)	Web 代理主机名或 IP 地址和端口
Username (用户名)	Web 代理用户名
Password (密码)	Web 代理密码

注册设备

第一次注册设备时，必须提供 Oracle 单点登录帐户。转到 [My Oracle Support \(http://support.oracle.com\)](http://support.oracle.com) 并单击“注册”以创建帐户。

▼ 使用 BUI 注册设备

1. 输入您的 Oracle 单点登录帐户用户名和密码。此时将显示 "Privacy Statement" (隐私声明)。可以随时在 BUI 和 CLI 中查看该声明。
2. 提交更改。
3. 使用 [My Oracle Support \(http://support.oracle.com/\)](http://support.oracle.com/) 可完成[自动服务请求 \(ASR\) \(http://oracle.com/asr\)](#) 的激活。请参见 "How To Manage and Approve Pending ASR Assets In My Oracle Support" (文档 ID 1329200.1)。

▼ 使用 CLI 注册设备

- 1. 将 `soa_id` 和 `soa_password` 分别设置为您的 Oracle 单点登录帐户的用户名和密码。
- 2. 提交更改。
- 3. 使用 [My Oracle Support \(http://support.oracle.com/\)](http://support.oracle.com/) 可完成[自动服务请求 \(ASR\) \(http://oracle.com/asr\)](http://oracle.com/asr) 的激活。请参见 "How To Manage and Approve Pending ASR Assets In My Oracle Support" (文档 ID 1329200.1) 。

例 11-1 CLI 注册

```
dory:> configuration services scrk
dory:configuration services scrk>set soa_id=myuser
soa_id = myuser(uncommitted)
dory:configuration services scrk> set soa_password=mypass
soa_password = ***** (uncommitted)
dory:configuration services scrk> commit
```

▼ 更改帐户信息

- 1. 单击 "Change account..." (更改帐户...) 可更改设备使用的 Oracle 单点登录帐户。
- 2. 提交更改。
- 3. 使用 My Oracle Support 可完成自动服务请求 (ASR) 的激活。请参见 "How To Manage and Approve Pending ASR Assets In My Oracle Support" (文档 ID 1329200.1)

回拨状态

表 11-45 回拨状态

属性	说明
Last heartbeat sent at (上次心跳发送时间)	上次向 Oracle 支持发送心跳的时间

回拨状态

如果在输入有效的 Oracle 单点登录帐户之前启用回拨服务，则该服务将显示为维护状态。必须输入有效的 Oracle 单点登录帐户才能使用回拨服务。

回拨日志

《Oracle ZFS Storage Appliance 客户服务手册 》中的“日志”中有回拨事件的日志。

REST

RESTful API

ZFSSA REST 风格的 API 允许您针对管理的资源 URL 路径，使用简单的请求（例如 GET、PUT、POST 和 DELETE HTTP）来管理 ZFSSA。

基于 ZFSSA RESTful 的架构定义为分层客户机 – 服务器模型。此模型的优势在于，在没有客户机配置的情况下，可以通过标准集线器、路由器及其他网络系统透明地重定向服务。此架构支持信息缓存，在许多客户机请求相同的静态资源时非常有用。

有关完整的 ZFSSA RESTful API 文档，请参见 Oracle ZFS Storage Appliance 文档。

服务标签

使用服务标签可以查询设备获取以下数据，从而为产品清单和支持提供便利：

- 系统序列号
- 系统类型
- 软件版本号

可以在 Oracle 支持系统中注册服务标签，从而轻松跟踪您的 Oracle 设备，同时加速服务呼叫的处理。默认情况下，服务标签处于启用状态。

服务标签属性

表 11-46 UDP/TCP 端口属性	
属性	说明
Discovery Port（搜索端口）	用于服务标签搜索的 UDP 端口。默认值是 6481
Listener Port（监听程序端口）	用于查询服务标签数据的 TCP 端口。默认值是 6481

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

SMTP 服务

SMTP 服务发送设备生成的所有邮件，通常是按“[Alerts](#)”（警报）屏幕中的配置对警报进行响应。SMTP 服务不接受外部邮件 – 它仅发送由设备自身自动生成的邮件。

默认情况下，SMTP 服务使用 DNS（MX 记录）确定将邮件发送到何处。如果未为设备的域配置 DNS，或者外发邮件的目标域未正确设置 DNS MX 记录，可以将设备配置为通过外发邮件服务器（通常称为智能主机）转发所有邮件。

SMTP 属性

表 11-47 SMTP 属性

属性	说明
Send mail through smarthost（通过智能主机发送邮件）	如果启用，将通过指定的外发邮件服务器发送所有邮件。否则，使用 DNS 确定对于特定域要将邮件发送到的位置。
Smarthost hostname（智能主机主机名）	外发邮件服务器主机名。
Allow customized from address（允许定制发件人地址）	如果启用，将电子邮件的 "From"（发件人）地址设置为 "Custom from address"（定制发件人地址）属性。例如，如果默认 "From"（发件人）地址被标识为垃圾邮件，则可以定制此项。
Custom from address（定制发件人地址）	用于外发电子邮件的 "From"（发件人）地址。

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

更改属性时，可以使用[“警报”](#)发送测试电子邮件来验证属性是否正确。未投递电子邮件的一个常见原因是 DNS 配置不正确，这会阻止设备确定要将邮件投递到的邮件服务器；如前面所述，如果无法配置 DNS，可以使用智能主机。

SMTP 日志

表 11-48 SMTP 日志

日志	说明
network-smtp:sendmail	记录 SMTP 服务事件
mail	SMTP 活动的日志（包括已发送的邮件）

SNMP 服务

SNMP (Simple Network Management Protocol, 简单网络管理协议) 服务在设备上提供两项不同的功能：

- SNMP 可以提供设备状态信息。
- 可以配置[第 9 章 警报配置](#)以发送 SNMP 陷阱。

启用此服务时可以使用 SNMP 版本 v1、v2c 和 v3。设备支持最多 50 个物理和逻辑网络接口。网络接口超过 50 个会导致此类命令 (例如 snmpwalk 和 snmpget) 超时。如果需要 50 个以上的网络接口, 请与 Oracle 支持联系。

SNMP 属性

- Version (版本) : v1/2c 和 v3 之间切换。
- Community name (团体名称) : 公共输入和用户输入之间切换。如果选择用户输入, 还必须输入团体名称。如果选择 v3, 则此属性不可用。
- Authorized network/subnet (授权网络/子网) : 输入适当的 IPv4 地址和子网 (0-32 的整数)。如果选择 v3, 则此属性不可用。
- Appliance contact (设备联系人) : 输入适当的设备联系人。
- Username/password (用户名/密码) : 输入有效的用户名 (最多 501 个字符) 和密码 (8-501 个字符)。如果选择 v1/2c, 则此属性不可用。
- Authentication (验证) : MD5 和 SHA 验证算法之间切换。如果选择 v1/2c, 则此属性不可用。
- Privacy (隐私) : None 和 DES 加密算法之间切换。如果选择 v1/2c, 则此属性不可用。
- Engine ID (引擎 ID) : EngineID 值经过 snmpd 的散列处理。如果此前未启用 SNMP, 标签将显示 "0x000"。
- Trap destinations (陷阱目的地) : 允许添加 IPv4 地址。使用 "+" 和 "-" 按钮添加或删除地址。

有关更改服务属性的信息, 请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

SNMP 服务还提供 MIB-II 位置字符串。此属性源于[“系统标识” \[258\]](#)配置。

SNMP MIB

如果 SNMP 服务处于联机状态, 授权网络将可以访问以下 MIB (Management Information Base, 管理信息库) :

表 11-49 SNMP MIB

MIB	用途
.1.3.6.1.2.1.1	MIB-II 系统 – 常规系统信息，包括主机名、联系人和位置
.1.3.6.1.2.1.2	MIB-II 接口 – 网络接口统计信息
.1.3.6.1.2.1.4	MIB-II IP – Internet 协议信息，包括 IP 地址和路由表
.1.3.6.1.4.1.42	Sun Enterprise MIB (SUN-MIB.mib.txt)
.1.3.6.1.4.1.42.2.195	Sun FM – 故障管理统计信息（下面链接的 MIB 文件）
.1.3.6.1.4.1.42.2.225	Sun AK – 设备信息和统计信息（下面链接的 MIB 文件）

注意：Sun MIB 文件位于 <https://your IP address or host name:215/docs/snmp/>

Sun FM MIB

Sun FM MIB (SUN-FM-MIB.mib) 提供对 SUN Fault Manager 信息的访问权限，例如：

- 系统上未解决的问题
- Fault Manager 事件
- Fault Manager 配置信息

要读取以下四个主表：

表 11-50 Sun FM MIB

OID	内容
.1.3.6.1.4.1.42.2.195.1.1	故障管理问题
.1.3.6.1.4.1.42.2.195.1.2	故障管理故障事件
.1.3.6.1.4.1.42.2.195.1.3	故障管理模块配置
.1.3.6.1.4.1.42.2.195.1.5	故障管理故障资源

有关完整说明，请参见上面链接的 MIB 文件。

Sun AK MIB

Sun AK MIB (SUN-AK-MIB.mib) 提供以下信息：

- 产品说明字符串和部件号
- 设备软件版本
- 设备和机箱序列号
- 安装、更新和引导时间
- 群集状态
- 共享资源状态 – 共享资源名称、大小、已用字节和可用字节

要读取以下三个主表：

表 11-51 Sun AK MIB

OID	内容
.1.3.6.1.4.1.42.2.225.1.4	常规设备信息
.1.3.6.1.4.1.42.2.225.1.5	群集状态
.1.3.6.1.4.1.42.2.225.1.6	共享资源状态

有关完整说明，请参见上面链接的 MIB 文件。

配置 SNMP

▼ 配置 SNMP 以提供设备状态

1. 设置团体名称、授权网络和联系人字符串。
2. 如果需要，将陷阱目标设置为远程 SNMP 主机，否则将此项设置为 127.0.0.1。
3. 应用/提交配置。
4. 重新启动该服务。

▼ 配置 SNMP 以发送陷阱

1. 设置团体名称、联系人字符串和陷阱目标。
2. 如果需要，将授权网络设置为允许 SNMP 客户机，否则将此项设置为 127.0.0.1/8。
3. 应用/提交配置。

4. 重新启动该服务。
5. 您必须配置警报以发送要接收的陷阱。有关警报的详细信息，请参见[第 9 章 警报配置](#)。

Syslog 服务

Syslog 转发服务在设备上提供两个不同的功能：

- 可以配置[第 9 章 警报配置](#)来向一个或多个远程系统发送 Syslog 消息。
- 设备上支持 syslog 的服务会将其 syslog 消息转发到远程系统。

syslog 消息是一条简短的事件消息，从设备传输到一个或多个远程系统（我们通常称为：intercontinental printf（洲际 printf））。消息包含以下元素：

- 设备，描述发出消息的系统组件的类型
- 严重性，描述与消息关联的状况的严重程度
- 时间戳，描述关联事件的时间（以 UTC 时间表示）
- 主机名，描述设备的规范名称
- 标签，描述发出消息的系统组件的名称有关消息格式的详细信息，请参见以下内容。
- 消息，描述事件本身。有关消息格式的详细信息，请参见以下内容。

syslog 接收器随大多数操作系统而提供，包括 Solaris 和 Linux。许多第三方和开源管理软件包也支持 syslog。syslog 接收器允许管理员将来自许多系统的消息聚合到单个管理系统并合并为单个日志文件集。

Syslog 转发可以配置为使用 RFC 3164 描述的“经典”输出格式，或者是 RFC 5424 描述的更新的版本化输出格式。Syslog 消息作为 UDP 数据报进行传输。因此，它们可能会被网络丢弃，如果发送系统内存不足或网络非常拥堵，则可能根本不发送这些消息。因此，管理员应该认识到在网络发生复杂故障的情况下，一些消息可能缺失或被丢弃。

Syslog 属性

表 11-52 Syslog 属性

属性	说明
Protocol Version（协议版本）	要使用的 Syslog 协议版本，经典版或现代版
Destinations（目标）	要将消息转发到的目标 IPv4 和 IPv6 地址的列表。

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

经典版 Syslog : RFC 3164

经典版 Syslog 协议包括编码为单个整数优先级的 facility 和 level 值以及 timestamp、hostname、tag 和消息主体。

tag 是以下所述标签之一。

hostname 是“系统标识” [258]配置定义的设备规范名称。

更新的 Syslog : RFC 5424

经典版 Syslog 协议包括编码为单个整数优先级的 facility 和 level 值以及版本字段 (1)、timestamp、hostname、tag、app-name 和消息主体。Sun Storage 系统转发的 syslog 消息将 RFC 5424 procid、msgid 和 structured-data 字段设置为 nil 值 (-)，指示这些字段不包含任何数据。

app-name 是以下所述标签之一。

hostname 是“系统标识” [258]配置定义的设备规范名称。

SYSLOG 消息格式

Syslog 协议自身不定义消息有效载荷的格式，而将其留给发送器来包括适当的任何种类的结构化数据或非结构化用户可读字符串。Sun Storage 设备使用 syslog 子系统标签 ak 来指示结构化的可解析消息有效载荷，下面将进行介绍。其他子系统标签指示任意用户可读文本，但是管理员应该考虑这些字符串格式不稳定并且可能会进行更改而不通知或者在 Sun Storage 软件的将来发行版中删除。

表 11-53 SYSLOG 消息格式

设备	标签名称	说明
daemon（守护进程）	ak	设备子系统的通用标签。所有警报都将标记 ak，指示后面为 SUNW-MSG-ID。
daemon（守护进程）	idmap	“身份映射” [233]服务用于 POSIX 和 Windows 身份转换。
daemon（守护进程）	smbd	“SMB 数据协议” [191]用于访问共享资源。

SYSLOG 警报消息格式

如果警报配置有 "Send Syslog Message" (发送 syslog 消息) 操作，它将生成 syslog 消息有效载荷，包含由以下标准字段组成的本地化文本。每个字段的前缀为字段名称（大写），后面是冒号和空格字符。

表 11-54 SYSLOG 警报消息格式

字段名	说明
SUNW-MSG-ID	与警报关联的稳定 Sun 故障消息标识符。将在 Sun 的故障消息目录中向生成管理员警报的每个系统状况和故障诊断分配一个永久的唯一标识符。这些标识符可以通过电话轻松读取或记在您的记事本中，它们链接到 sun.com/msg/ 中的相应知识文章。
TYPE (类型)	状况的类型。这将是以下标签之一："Fault" (故障)，指示硬件组件或连接器故障；"Defect" (缺陷)，指示硬件缺陷或错误配置；"Alert" (警报)，指示未与故障或缺陷关联的状况，例如备份活动或远程复制完成。
VER	此编码格式自身的版本。此说明与 SUNW-MSG-ID 格式的版本 "1" 相对应。如果 VER 字段中为 "1"，解析代码可能会认为所有后面的字段都将存在。如果指定大于 1 的十进制整数，应该编写解析代码来处理或忽略其他字段。
SEVERITY (严重性)	与触发警报的问题相关联的状况的严重程度。严重程度列表如下所示。
EVENT-TIME (事件时间)	与此事件对应的时间。时间将为 "Day Mon DD HH:MM:SS YYYY" 格式的 UTC 时间。例如：Fri Aug 14 21:34:22 2009。
PLATFORM (平台)	设备的平台标识符。此字段仅供 Oracle 服务人员使用。
CSN	设备的机箱序列号。
HOSTNAME (主机名)	“系统标识” [258] 配置所定义的设备的规范名称。
SOURCE (源)	触发事件的设备软件内的子系统。此字段仅供 Oracle 服务人员使用。
REV	子系统的内部修订版。此字段仅供 Oracle 服务人员使用。
EVENT-ID (事件 ID)	与此事件关联的通用唯一标识符 (Universally Unique Identifier, UUID)。Oracle 的故障管理系统将 UUID 与每个警报和故障诊断关联，从而使管理员可以收集与单个状况关联的多个消息并使其相关，同时检测重复消息。Oracle 服务人员可以使用 EVENT-ID 检索与问题关联的、可能帮助 Oracle 对问题作出响应的其他事后信息。
DESC	与事件关联的状况的说明。

字段名	说明
AUTO-RESPONSE (自动响应)	系统中包括的故障管理软件对问题（如果有）的自动响应。自动响应包括主动使故障磁盘、DRAM 内存芯片以及处理器内核脱机等功能。
REC-ACTION (维修操作)	建议的维修操作。这将包括建议操作的简要摘要，但是管理员应该查看知识文章和本文档，以获得关于完整修复过程的信息。

SEVERITY (严重性) 字段将设置为下面一个值：

表 11-55 SYSLOG 严重性字段

严重性	Syslog 级别	说明
Minor (次要)	LOG_WARNING	发生当前不影响服务的状况，但是需要在该状况变得更严重之前对其进行更正。
Major (重大)	LOG_ERR	发生影响服务但是不严重的状况。
Critical (严重)	LOG_CRIT	发生严重影响服务并需要立即更正的状况。

接收器配置示例

大多数操作系统包括 syslog 接收器，但是可能需要执行一些配置步骤来开启该接收器。常见操作系统的一些示例如下所示。查看您的操作系统或管理软件的文档，以获得 syslog 接收器配置的特定详细信息。

配置 Solaris 接收器

Solaris 包括捆绑的 syslogd (1M)，其可以用作 syslog 接收器，但是默认情况下禁用远程接收功能。要允许 Solaris 接收 syslog 通信，请使用 svccfg 和 svcadm 修改 syslog 设置，如下所示：

```
# svccfg -s system/system-log setprop config/log_from_remote = true
# svcadm refresh system/system-log
```

Solaris syslogd 仅理解经典版 Syslog 协议。有关如何配置已接收消息的过滤和记录的信息，请参阅 syslog.conf(4) 手册页。

默认情况下，Solaris syslogd 将消息记录到 /var/adm/messages 中，测试警报将记录为如下所示：

```
Aug 14 21:34:22 poptart.sf.fishpong.com poptart ak: SUNW-MSG-ID: AK-8000-LM, \
TYPE: alert, VER: 1, SEVERITY: Minor\nEVENT-TIME: Fri Aug 14 21:34:22 2009\n\
PLATFORM: i86pc, CSN: 12345678, HOSTNAME: poptart\n\
SOURCE: jsui.359, REV: 1.0\n\
EVENT-ID: 92dfcb39-6e15-e2d5-a7d9-dc3e221becea\n\
DESC: A test alert has been posted.\n\
AUTO-RESPONSE: None.\nIMPACT: None.\nREC-ACTION: None.
```

配置 Linux 接收器

大多数 Linux 分发版都包括捆绑的 syslogd(8) 守护进程，其可以用作 syslog 接收器，但是默认情况下禁用远程接收功能。要允许 Linux 接收 syslog 通信，请编辑 /etc/sysconfig/syslog 配置文件，从而包括 -r 选项（启用远程日志记录）：

```
SYSLOGD_OPTIONS="-r -m 0"
```

然后重新启动日志记录服务：

```
# /etc/init.d/syslog stop
# /etc/init.d/syslog start
```

一些 Linux 分发版具有 ipfilter 包过滤器，它在默认情况下将拒绝 syslog UDP 包，必须修改该过滤器以允许这些包。在这些分发版上，使用与下面类似的命令添加 INPUT 规则以接受 syslog UDP 包：

```
# iptables -I INPUT 1 -p udp --sport 514 --dport 514 -j ACCEPT
```

默认情况下，Linux syslogd 将消息记录到 /var/log/messages 中，测试警报将记录为如下所示：

```
Aug 12 22:03:15 192.168.1.105 poptart ak: SUNW-MSG-ID: AK-8000-LM, \
TYPE: alert, VER: 1, SEVERITY: Minor EVENT-TIME: Wed Aug 12 22:03:14 2009 \
PLATFORM: i86pc, CSN: 12345678, HOSTNAME: poptart SOURCE: jsui.3775, REV: 1.0 \
EVENT-ID: 9d40db07-8078-4b21-e64e-86e5cac90912 \
DESC: A test alert has been posted. AUTO-RESPONSE: None. IMPACT: None. \
REC-ACTION: None.
```

系统标识

此服务提供系统名称和位置的配置。如果设备移至不同网络位置或改变用途，可能需要更改这些配置。

系统标识属性

表 11-56 系统标识属性

属性	说明
System Name (系统名称)	显示在用户界面中的设备的单个规范化标识名称。此名称独立于用于连接到系统的任何 DNS 名称（其将在远程 DNS 服务器上配置）。可以随时更改该名称
System Location (系统位置)	描述设备实际位置的文本字符串。如果已启用“SNMP” [251]，该项将作为 MIB-II 中的 <i>syslocation</i> 字符串导出

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

系统标识日志

表 11-57 系统标识日志

日志	说明
system-identity:node	记录系统标识服务事件和错误

SSH 服务

SSH (Secure Shell，安全 Shell) 服务允许用户登录到设备 CLI 并执行可以在 BUI 中执行的大多数相同管理操作。SSH 服务还可以用作从远程主机执行自动脚本的方式，例如用于检索每日日志或[《Oracle ZFS Storage Appliance Analytics 指南》中的“Analytics”统计信息](#)。

SSH 属性

表 11-58 SSH 属性

属性	说明	示例
Server key length (服务器密钥长度)	临时密钥中的位数。	768
Key regeneration interval (密钥重新生成时间间隔)	临时密钥重新生成时间间隔，以秒为单位。	3600

属性	说明	示例
Login grace period（登录宽限期）	如果客户机验证失败，SSH 连接会在此秒数后断开连接。	120
Permit root login（允许 root 用户登录）	允许 root 用户使用 SSH 登录。	有

有关更改服务属性的信息，请参见[“使用 BUI 配置服务” \[179\]](#)和[“使用 CLI 配置服务” \[181\]](#)。CLI 属性名称是上面列出的名称的短版本。

SSH 日志

表 11-59 SSH 日志

日志	说明
network-ssh:default	SSH 服务事件和错误的日志

配置 SSH

▼ 禁用 root SSH 访问权限

1. 将 "Permit root login"（允许 root 用户登录）设置为 false。
2. 应用/提交配置。

◆◆◆ 第 12 章

共享资源、项目和模式

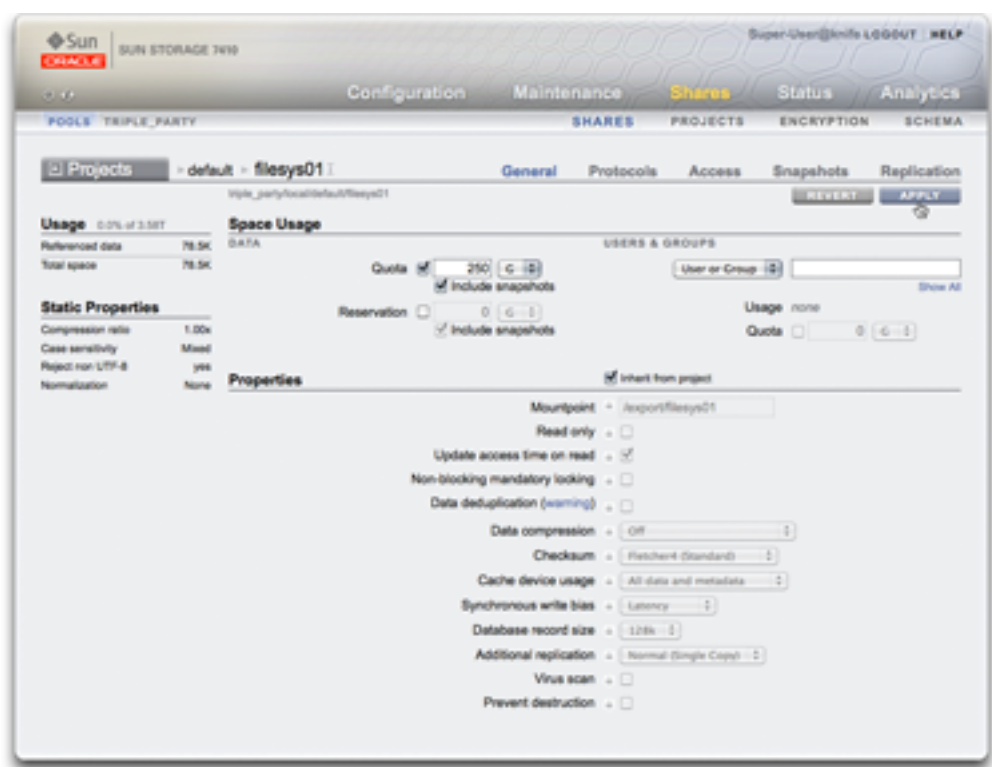
本节介绍 ZFSSA 共享资源、项目和模式。

用于常见管理目的，包括空间管理和常用设置，共享资源可以组合为 ZFSSA 项目。除了标准的内置属性之外，还可以配置可在所有共享资源和项目上使用的任意数目的其他属性。可以为这些属性指定基本类型以用于验证目的，并且它们可以与大多数其他标准属性一样被继承。这些值从不会被软件以任何方式使用，仅供最终用户使用。属性模式对系统来说是全局性的，在所有池之间保持一致，并且在群集对等设备之间保持同步。

了解共享资源

存储池

图 12-1 相似的共享资源可分组到一个项目中。



ZFSSA 基于 ZFS 文件系统。ZFS 将底层存储设备组合到池中，并根据需要从此存储分配文件系统和 LUN。创建文件系统或 LUN 之前，必须首先在 ZFSSA 上[第 5 章 存储配置](#)。配置存储池后，不需要静态调整文件系统大小，尽管可以使用[“共享资源空间管理” \[265\]](#)实现此行为。

支持多个存储池时，通常不建议使用这种类型的配置，因为它具有明显的缺点，如[第 5 章 存储配置](#)一节中所述。仅当两个不同配置文件的性能或可靠性特性区别很大时，才应该使用多个池，例如一个镜像池用于数据库，一个 RAID-Z 池用于流化处理工作负荷。

当多个池在单个主机上处于活动状态时，BUI 将在菜单栏中显示一个下拉列表，可用于在池之间进行切换。在 CLI 中，当前池的名称将显示在括号中，并且可以通过设置 'pool' 属性来更改该池。如果仅配置了一个池，则将隐藏这些控件。选择多个池时，UI 所选择的默认池是任意的，所以任何脚本操作都应该确保在处理任何共享资源之前显式设置池名称。

使用共享资源

共享资源是通过支持的数据协议导出到 ZFSSA 的客户机的文件系统和 LUN。文件系统导出基于文件的分层结构，可以通过“SMB” [191]、“NFS” [185]、“HTTP/WebDav” [208]和“FTP” [206]访问。LUN 导出基于块的卷，可以通过“iSCSI” [189]或光纤通道访问。项目/共享资源元组是池内共享资源的唯一标识符。多个项目可以包含具有相同名称的共享资源，但是单个项目不能包含具有相同名称的共享资源。单个项目可以包含文件系统和 LUN，它们共享相同名称空间。

共享资源属性

所有项目和共享资源都有许多关联的属性。这些属性分为以下组：

表 12-1 项目和共享资源属性

属性类型	说明
继承	这是最常见的属性类型，表示大多数可配置项目和共享资源属性。属于项目一部分的共享资源可以具有属性的本地设置，它们还可以从父项目继承其设置。默认情况下，共享资源从项目继承所有属性。如果更改项目上的属性，将更新继承该属性的所有共享资源以反映新值。继承后，所有属性具有与父项目相同的值，挂载点和 SMB 属性除外。继承后，这些属性将项目设置与其自己的共享资源名称串联。
只读	这些属性表示项目和共享资源的统计信息，无法更改。这种类型的最常见属性是空间使用情况统计信息。
空间管理	这些属性（配额和预留空间）应用于共享资源和项目，但是无法继承。将在所有共享资源中强制配额为 100G 的项目，但是每个单独共享资源将没有配额，除非明确设置。
创建时	可以在创建文件系统或 LUN 时指定这些属性，但是一旦创建了共享资源则无法更改这些属性。这些属性控制盘上数据结构，包括国际化设置、区分大小写以及卷块大小。
项目默认	在项目上设置这些属性，但对项目自身没有影响。创建文件系统或 LUN 时可以使用它们填充初始设置，

属性类型	说明
	并且共享资源具有通用非可继承属性集时这些属性非常有用。更改这些属性不会影响现有共享资源，可以在创建共享资源之前或之后更改这些属性。
文件系统本地	这些属性仅应用于文件系统，可方便管理文件系统的根目录。无法在项目上设置它们。还可以通过带内协议操作设置这些访问控制属性。
LUN 本地	这些属性仅适用于 LUN，无法继承。无法在项目上设置它们。
定制	这些是用户定义的属性。有关更多信息，请参见 “模式” [322] 。

共享资源快照

快照是文件系统或 LUN 的即时副本。可以手动或通过设置自动调度表来创建快照。快照最初不占用额外空间，但是随着活动共享资源更改，将保留以前未引用的块作为最后快照的一部分。逐渐地，最后快照将占用额外空间，最大大小等于创建快照时的文件系统大小。

可以通过标准协议在文件系统根目录的 `.zfs/snapshot` 快照中访问文件系统快照。默认情况下隐藏该目录，仅能通过显式更改为 `.zfs` 目录来访问该目录。可以在[“快照” \[303\]](#)视图中更改此行为，但是可能会导致备份软件除了备份实时数据之外还备份快照。LUN 快照无法直接访问，但它们可以用作回滚目标或克隆源。项目快照包括项目内所有共享资源的快照，通过名称标识快照。如果对属于较大项目快照一部分的共享资源快照重命名，它将不再被视为该同一快照的一部分，如果任何快照重命名为与父项目中的快照具有相同名称，则会将其视为该父快照的一部分。

共享资源支持回滚到先前快照的功能。发生回滚时，将销毁所有较新快照（以及较新快照的克隆），活动数据将恢复到创建该快照时的状态。快照仅包括数据，不包括属性，因此自创建该快照后更改的任何属性设置将保持不变。

共享资源克隆

许可声明：可以免费试用远程复制和克隆特性，但若要在生产环境中使用，必须为各特性购买单独的许可证。在试用期后，必须为这些特性获取许可证或停用它们。*Oracle 保留随时审计许可合规性的权利。有关详细信息，请参阅 “Oracle Software License Agreement (“SLA”) and Entitlement for Hardware Systems with Integrated Software Options”。*

克隆是共享资源快照的可写副本，视为用于管理目的的独立共享资源。与快照类似，克隆最初不占用额外空间，但是随着新数据写入克隆，新更改所需的空间将与该克隆关

联。不支持项目克隆。因为在快照与克隆之间共享空间，并且快照可以具有多个克隆，所以销毁快照也会销毁活动克隆。

共享资源空间管理

在 7000 系列中，文件系统和 LUN 在管理物理存储方面的行为与在许多其他系统上有所不同。如“概念” [262] 页中所述，该 ZFSSA 使用一个池存储模型，在该模型中，所有文件系统和 LUN 都共享公用的空间。文件系统从不会向它们分配显式大小，并且只根据需求占用相应的空间量。LUN 会预留足够的物理空间来写入设备的完整内容，除非它们是瘦置备的，在这种情况下，LUN 的行为像文件系统一样，并且仅使用由数据以物理方式占用的空间量。

在一般情况下信任用户会执行正确操作的环境中，此系统可以提供最大的灵活性并最大程度地简化了管理。较严格的环境（在这种环境中，用户对数据的使用受监视和/或限制）需要更仔细的管理。本节介绍了可供管理员用来控制和管理空间使用情况的一些工具。

共享资源空间术语

在开始详细介绍之前，需要了解一些在谈论 ZFSSA 上的空间使用情况时所使用的基本术语，这一点很重要。

- **Physical Data（物理数据）** – 在磁盘上以物理方式存储的数据的大小。通常，这等同于对应数据的逻辑大小，但在压缩阶段中或由于其他因素，可能会有所不同。这包括活动共享资源以及所有快照的空间。空间计算通常是基于物理空间执行和管理的。
- **Logical Data（逻辑数据）** – 文件系统占用的逻辑空间量。这没有考虑压缩因素，并且可以将其视为文件系统占用的空间量的理论上界。使用不同的压缩算法将文件系统复制到其他 ZFSSA 不会占用比这更多的空间量。此统计信息未显式导出，通常只能通过获取所占用的物理空间量并将其与当前的压缩比率相乘来计算获得。
- **Referenced Data（引用数据）** – 这表示活动共享资源所引用的空间总量，且独立于任何快照。这是销毁所有快照后共享资源将占用的空间量。这同时也是用户可以通过数据协议直接管理的数据量。
- **Snapshot Data（快照数据）** – 这表示共享资源的所有快照当前保存的数据总量。这是销毁所有快照时将释放的空间量。
- **Quota（配额）** – 配额表示对可以供任何特定实体占用的空间量施加的限制。它可以基于文件系统、项目、用户或组，并且独立于任何当前空间使用情况。
- **Reservation（预留空间）** – 预留空间表示保证可以供特定项目或文件系统使用的空间。这将从池的剩余部分获取可用空间，不会增加文件系统占用的实际空间。此设置无法应用于用户和组。通过将配额和预留空间设置为相同的值可以创建传统意义上的静态大小的文件系统。

了解快照

快照给空间管理带来了一个进退两难的选择。它们表示在给定的时间点由某个共享资源引用的一组物理块。最初，此快照不占用额外空间。但是随着在新共享资源中覆盖新数据，活动共享资源中的块将只包含新数据，较旧的块将“保存”在最新（可能较旧）的快照中。逐渐地，随着活动共享资源中内容的不断分化，快照会占用更多的空间。

某些其他系统试图隐瞒快照的成本，因而谎称使用快照是免费的，或者“预留”空间专门存放快照数据。这些系统试图掩盖快照所固有的基本事实。如果创建了任何给定规模的文件系统的快照，并在文件系统中重写全部数据，很显然必须维护对两倍于文件系统原有数据的数据的引用。使用快照并不是免费的，其他系统要让这种不切实际的说法得以成立，唯一的办法只能在空间占满时悄悄地销毁快照。但这种做法通常是极其恶劣的，因为一个进程不受控制地胡乱重写数据可能导致所有先前的快照都被销毁，从而阻碍任何进行中的恢复操作。

在 Sun Storage 7000 系列中，快照成本始终是明确可见的，同时提供了相关工具，可以按照最为符合特定环境的管理模型的方式来管理快照空间。每个快照都有两个关联的空间统计信息：唯一空间和引用空间。引用空间量是文件系统在创建快照时所占用的总空间量。它表示快照在理论上的最大大小，前提是该快照仍然是所有数据块的唯一引用。唯一空间指的是只由当前快照所引用的物理空间。当快照销毁时，其唯一空间可由池中的其他使用者使用。请注意，所有快照所使用的空间总量并不等于所有快照的唯一空间之和。当有一个共享资源和一个快照时，所有块必须由该共享资源和该快照其中之一或由这两者引用。但是，当有多个快照时，一个块可由快照的某个子集引用，而不是由某个快照引用。例如，如果在创建文件时创建了快照 X 和 Y，在删除文件时创建了另一个快照 Z，则文件中的块将保存在 X 和 Y 中，而不会保存在 Z 中。在这种情况下，销毁 Z 不能释放空间，但同时销毁 X 和 Y 则可以。因此，销毁任何快照都可能影响相邻快照所引用的唯一空间，但是由快照占用的总空间量始终会不断减少。

项目或共享资源的总大小始终能够反映出所有快照占用的空间总量，但是也可以知道快照各自细分的空间使用量。在项目级别可以设置配额和预留空间，对此总空间强制应用物理约束。此外，还可以在文件系统级别设置配额和预留空间，这些设置只能应用于引用的数据或总数据。是否应当向引用的数据或总的物理数据应用配额和预留空间取决于管理环境。如果用户未控制其快照（例如，为其设置自动快照日程表），则在配额计算中通常不应包括快照。否则，用户可能会耗尽空间，但在无法删除文件时会感到困惑。如果不了解快照或用于管理这些快照的方式，则在没有管理员干预的情况下可能无法解决此类情况。在这种情况下，快照是应当计入用于提供备份功能的系统操作的开销成本。另一方面，在某些环境中，将根据用户的物理空间需求为其记帐，在这种情况下，快照表示用户的一种选择，用以根据其数据集的变动率提供一定级别的备份来满足其需求。在这些环境中下，根据总的物理数据（包括快照）实施配额才更合理。用户了解快照的成本，并且可以为用户提供一些方法（例如通过 ZFSSA 上的专用角色）来主动管理快照。

文件系统和项目设置

实施配额和预留空间的最简单方式是以每个项目或每个文件系统为单位。配额和预留空间不应用于 LUN，尽管对它们的使用被计入了总的项目配额或预留空间。

数据配额

数据配额对文件系统或项目可使用的空间量施加限制。默认情况下，它将包括文件系统以及所有快照中的数据。当文件系统已满时，尝试写入新数据的客户机会因为配额或因为存储池的空间不足而收到一个错误。如[“快照” \[265\]](#)一节中所述，此行为并非在所有情况下都很直观，尤其是当存在快照时。如果数据块被快照所引用，则删除文件可能会导致文件系统写入新数据，因此，在这种情况下，销毁现有快照可能是降低空间使用率的唯一方法。

如果未设置 "include snapshots"（包含快照）属性，则配额仅应用于文件系统引用的即时数据，不应用于任何快照。快照使用的空间受项目级别的配额限制，但不受其他限制。在这种情况下，删除快照引用的文件将导致文件系统的被引用数据减少，尽管系统作为整体正使用更多的空间。如果存储池已满（与文件系统达到预设的配额相对），则销毁快照可能是释放空间的唯一方法。

数据配额是严格实施的，这意味着在空间使用率接近限制时，必须对可写入的数据的量进行限制，因为只有在确认了写入操作后才能获知要写入的精确数据量。当在已达到配额或将要达到配额的情况下运行时，这可能会影响性能。因此，在正常运行过程中，保持在配额以下通常是明智的。

配额通过 BUI 在 "Shares"（共享资源）> "General"（常规）> "Space Usage"（空间使用情况）> "Data"（数据）下进行管理。它们在 CLI 中作为 quota 和 quota_snap 属性进行管理。

数据预留空间

数据预留空间用于确保文件或项目至少有一定量的可用空间，即使文件系统或其他共享资源尝试使用更多的空间。此未使用的预留空间被视为文件系统的一部分，因此，如果池（或项目）的其余部分已达到容量，该文件系统仍可以写入新数据，即使其他共享资源可能已经空间不足。

默认情况下，预留空间包括文件系统的所有快照。如果未设置 "include snapshots"（包含快照）属性，则预留空间仅应用于文件系统的即时数据。如[“快照” \[265\]](#)一节所述，创建快照时的行为可能并不总是直观的。如果实施了用于文件系统数据（但不是快照）的预留空间，则每次创建快照时，系统都必须预留足够的空间以便该快照完全分化，即使从不发生此操作。例如，如果没有快照时，一个 50G 的文件系统有 100G 的预留空间，则在创建第一个快照时将预留 50G 的额外空间，该文件系统最终将总共预留 150G 空间。如果没有足够的空间来确保数据的完全分化，则创建快照将失败。

预留空间通过 BUI 在 "Shares" (共享资源) > "General" (常规) > "Space Usage" (空间使用情况) > "Data" (数据) 下进行管理。它们在 CLI 中作为 reservation 和 reservation_snap 属性进行管理。

复制 LUN 时的空间管理

创建 LUN 时，系统会预留您为 LUN 配置的全部物理空间，并且这些空间无法被其他文件系统使用（除非进行了瘦置备）。对于复制，如果创建了任何给定规模的 LUN 的快照，则还会预留最多两倍于 LUN 的大小，具体取决于使用的 LUN 空间。

以下列表显示复制 LUN 时所需的最大开销空间：

- 在更新之间在源上最多 100%
- 更新过程中在源上最多 200%
- 在目标上最多 200%

用户和组设置

查看当前使用量

不管是否使用了用户和组配额，都可以针对文件系统和项目以用户或组为单位查询当前使用量。在使用此功能之前，在此软件的较旧版本上创建的存储池可能需要应用《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“更新”。在应用延迟的更新后，可能需要花费一些时间才能将所有文件系统升级到支持以用户和组为单位控制空间使用和配额的版本。

▼ 在 BUI 中查看当前使用量

1. 要在 BUI 中查看当前使用量，请转至 "Shares" (共享资源) > "Shares" (共享资源) > "General" (常规)。
2. 在 "Space Usage" (空间使用情况) - "Users and Groups" (用户和组) 一节中单击 "User" (用户) 或 "Group" (组) 下拉菜单以选择 "User" (用户) 或 "Group" (组)，并查询共享资源中或项目中任何给定用户或组的当前使用量。
3. 键入您要查询的 "User" (用户) 或 "Group" (组) 的名称。键入后查询即进行处理。

查询完成后，将显示当前使用量。此外，"Show All" (全部显示) 链接将显示一个对话框，其中显示了所有用户或组的当前空间使用量列表。此对话框只能查询一个特定类型（用户或组），不支持同时查询这两者。此列表显示规范的 UNIX 和 Windows 名称（如果已启用映射）以及使用量和（对于文件系统）配额。

▼ 在 CLI 中查看当前使用量

1. 在 CLI 中，请在特定项目或共享资源的上下文中使用 **users** 和 **groups** 命令。
2. 使用 **show** 命令以表格形式显示当前使用量。
3. 要检索特定用户或组的使用量，请选择所需的用户或组并使用 **get** 命令。

```
clownfish:> shares select default
clownfish:shares default> users
clownfish:shares default users> list
USER      NAME                USAGE
user-000   root                 325K
user-001   ahl                  9.94K
user-002   eschrock             20.0G
clownfish:shares default users> select name=eschrock
clownfish:shares default user-002> get
      name = eschrock
  unixname = eschrock
    unixid = 132651
   winname = (unset)
    winid  = (unset)
     usage = 20.0G
```

设置用户或组配额

可以在文件系统级别为用户或组设置配额。这些配额将基于文件或目录的所有者或组的 POSIX 或 Windows 身份来限制物理数据使用量。在用户和组配额与文件系统和项目数据配额之间存在一些显著差异：

- 用户和组配额只能应用于文件系统。
- 用户和组配额是使用延迟的执行实施的。这意味着在数据写入到磁盘之前，用户将能够在短时间内超出其配额。在数据被推送到磁盘后，用户将收到一个有关新写入操作的错误，与文件系统级别的配额情况类似。
- 用户和组配额始终是针对引用的数据实施的。这意味着快照不影响任何配额，快照的克隆将占用与实施的配额相同的空间量，即使底层块是共享的。
- 不支持用户和组预留空间。
- 与数据配额不同，用户和组配额与常规文件系统数据存储在一起。这意味着，如果文件系统空间不足，您将无法更改用户和组配额。必须首先使更多的空间可用，才能修改用户和组配额。
- 在执行远程复制期间，将发送用户和组配额。管理员应当确保源和目标上的名称服务环境相同。
- 整个共享资源的 NDMP 备份和恢复将包括任何用户和组配额。恢复到现有共享资源将不会影响任何现有配额。

▼ 使用 BUI 设置用户或组配额

1. 在 BUI 中，转至 "Shares" (共享资源) > "Shares" (共享资源) > "General" (常规)。
2. 在 "Space Usage" (空间使用情况) – "Users and Groups" (用户和组) 一节中单击 "User" (用户) 或 "Group" (组) 下拉菜单以选择 "User" (用户) 或 "Group" (组)，并查询共享资源中或项目中任何给定用户或组的当前使用量。
3. 在浏览器中，可以从["general" \(常规\) \[284\]](#)选项卡中的 "Space Usage" (空间使用情况) -> "Users & Groups" (用户和组) 下管理用户配额。与查看使用量一样，在您键入用户或组时将显示当前使用量。完成用户名或组名的输入并显示当前使用量后，可通过选中 "quota" (配额) 旁边的框并在 "size" (大小) 字段中输入一个值来设置配额。要禁用配额，请取消选中此框。在进行任何更改后，单击 "Apply" (应用) 按钮应用更改。
4. 虽然页面上的所有属性是一起提交的，但是用户和组配额将与其他属性分开验证。如果输入了一个无效的用户和组以及另一个无效的属性，则只能显示其中的一个验证错误。在更正该错误后，再次尝试应用更改时将显示另一个错误。

▼ 使用 CLI 设置用户或组配额

- 在 CLI 中，可从共享资源上下文中使用 'users' 或 'groups' 命令管理用户配额。可以通过选择特定用户或组并使用 "set quota" 命令来设置配额。未占用文件系统上的任何空间且未设置任何配额的任何用户将不会显示在活动用户列表中。要为这样的用户或组设置配额，请先使用 "quota" 命令，之后可以设置名称和配额。要清除某个配额，请将其值设置为 "0"。

```
clownfish:> shares select default select eschrock
clownfish:shares default/eschrock> users
clownfish:shares default/eschrock users> list
USER      NAME      USAGE  QUOTA
user-000   root      321K   -
user-001   ahl       9.94K   -
user-002   eschrock  20.0G   -
clownfish:shares default/eschrock users> select name=eschrock
clownfish:shares default/eschrock user-002> get
      name = eschrock
      unixname = eschrock
      unixid = 132651
      winname = (unset)
      winid = (unset)
      usage = 20.0G
      quota = (unset)
clownfish:shares default/eschrock user-002> set quota=100G
      quota = 100G (uncommitted)
clownfish:shares default/eschrock user-002> commit
clownfish:shares default/eschrock user-002> done
clownfish:shares default/eschrock users> quota
clownfish:shares default/eschrock users quota (uncommitted)> set name=bmc
```

```

name = bmc (uncommitted)
clownfish:shares default/eschrock users quota (uncommitted)> set quota=200G
quota = 200G (uncommitted)
clownfish:shares default/eschrock users quota (uncommitted)> commit
clownfish:shares default/eschrock users> list

```

USER	NAME	USAGE	QUOTA
user-000	root	321K	-
user-001	ahl	9.94K	-
user-002	eschrock	20.0G	100G
user-003	bmc	-	200G

身份管理

用户和组配额利用 ZFSSA 上的“身份映射” [233] 服务。这允许将用户和组指定为 UNIX 或 Windows 身份，具体取决于环境。与文件所有权一样，将采用以下方式跟踪这些身份：

- 如果没有 UNIX 映射，将存储对 Windows ID 的引用。
- 如果存在 UNIX 映射，则存储 UNIX ID。

这意味着身份的规范形式是 UNIX ID。如果以后更改了映射，则会基于新的 UNIX ID 实施新映射。如果 Windows 用户创建某个文件时不存在映射，并且之后创建了映射，则出于访问控制和使用格式目的，会将新文件视为拥有一个不同的所有者。这也意味着如果重用某个用户 ID（例如创建新的用户名关联），则任何现有文件或配额将显示为由新用户名拥有。

我们建议在尝试主动使用文件系统之前，请先建立任何身份映射规则。否则，映射中的任何更改有时会产生意外的结果。

文件系统名称空间

必须为 ZFSSA 上的每个文件系统提供唯一的挂载点来充当文件系统数据的访问点。可以为项目提供挂载点，但是这些挂载点仅作为一个工具使用继承的属性来管理名称空间。项目从不挂载，且不通过任何协议导出数据。

所有共享资源必须挂载在 /export 下。虽然可以创建挂载在 /export 下的文件系统，但是没有必要。如果不存在这样的共享资源，则将根据需要在此层次结构部分中动态创建任意目录。每个挂载点在群集中必须是唯一的。

嵌套了挂载点的名称空间

可以在一个文件系统下创建带有挂载点的其他文件系统。在这种情况下，应先挂载父文件系统，再挂载子文件系统，反之亦然。使用嵌套的挂载点时应考虑以下事项：

- 如果挂载点不存在，将创建一个挂载点，由 root 和模式 0755 所拥有。在重命名、销毁或移动文件系统时，此挂载点可能会也可能不会被卸载，视情况而定。为安全起见，应在创建子文件系统之前在父共享资源中创建挂载点。
- 如果父目录是只读的，且挂载点不存在，则文件系统挂载将失败。这种情况可能会在创建文件系统时同步发生，但也可能会在进行大规模更改（例如使用继承挂载点重命名文件系统）时异步发生。
- 重命名文件系统或更改其挂载点时，当前挂载点下的所有子挂载点以及新挂载点（如果不同）将被卸载并在应用更改后被重新挂载。这将中断当前访问共享资源的任何数据服务。
- 如下面所概述，是否支持自动遍历嵌套挂载点取决于协议。

挂载点的协议访问名称空间

无论协议设置如何，每个文件系统都必须有一个挂载点。但是，使用这些挂载点的方式取决于协议。

NFSv2 / NFSv3 名称空间

在 NFS 下，每个文件系统是可通过 MOUNT 协议设为可视的一项独特导出内容。NFSv2 和 NFSv3 无法遍历嵌套文件系统，必须通过文件系统的完整路径来访问每个文件系统。虽然嵌套挂载点仍可起作用，但是尝试遍历嵌套挂载点将会导致客户机上显示空目录。虽然通过使用自动挂载方式进行挂载可以缓解这种现象，但是在动态环境中透明支持嵌套挂载点则需要使用 NFSv4。

NFSv4 名称空间

相较于 NFSv3，NFSv4 在处理挂载点时进行了多项改进。首先是，可以挂载父目录，即使分层结构中该挂载点没有可用共享资源也是如此。例如，如果共享了 /export/home，则可以在客户机上挂载 /export，并透明地遍历到实际的导出内容。更为重要的是，某些 NFSv4 客户机（包括 Linux）支持自动客户端挂载，有时称为“镜像挂载”。使用这样的客户机，当用户遍历某个挂载点时，子文件系统将自动挂载到相应的本地挂载点，并且当文件系统在客户机上被卸载时也同时卸载。从服务器的角度来看，这些是单独的挂载请求，但在客户机中它们缝合在一起以形成无缝的文件系统名称空间。

SMB 名称空间

SMB 协议不使用挂载点，因为通过资源名称来使每个共享资源变为可用。但是，每个文件系统必须仍具有唯一挂载点。当前不支持嵌套挂载点（一个资源内有多个文件系统），任何遍历挂载点的尝试将会导致显示空目录。

FTP/FTPS/SFTP 名称空间

使用文件系统的标准挂载点导出文件系统。完全支持嵌套挂载点，并且这些挂载点对于用户透明。但是，不能在嵌套文件系统的父文件系统处于共享状态时不共享嵌套文件系统。如果共享父挂载点，则也将共享所有子挂载点。

HTTP/HTTPS 名称空间

文件系统导出在 /shares 目录下，因此 /export/home 下的文件系统将通过 HTTP/HTTPS 显示在 /shares/export/home 下。完全支持嵌套挂载点，并且这些挂载点对于用户透明。FTP 协议部分所述的有关共享资源选项冲突的相同行为也适用于 HTTP。

Shares (共享资源) > Shares (共享资源)

在 BUI 中使用 "Shares" (共享资源) > "Shares" (共享资源)

通过 "Shares" (共享资源) > "Shares" (共享资源) 可访问 "Shares" (共享资源) UI。默认视图显示了系统上所有项目间的共享资源。

共享资源列表

默认视图列出了系统上的所有共享资源。在此列表中，您可以重命名共享资源、在项目之间移动共享资源，以及编辑各个共享资源。共享资源分为两个列表，即 "Filesystems" (文件系统) 和 "LUNs" (LUN)，可以通过在此视图中切换选项卡来选择这两个列表。每个共享资源均包含以下字段：

表 12-2 BUI 共享资源列表

字段	说明
Name (名称)	共享资源的名称。如果查看所有项目，也会包括项目名称。共享资源名称是可编辑的文本字段。通过单击名称，可以输入一个新名称。按回车键或从名称移动焦点将会提交更改。系统将要求您确认该项操作，因为对共享资源进行重命名需要断开活动客户机的连接。
Size (大小)	对于文件系统，这是文件系统的总大小。对于 LUN，它是卷（可能进行了瘦置备也可能没有进行瘦置备）的大小。有关更多信息，请参见 "使用情况统计信息" [274] 。

字段	说明
Mountpoint (挂载点)	文件系统的挂载点。这是通过 NFS 提供的路径，对于 FTP 和 HTTP 为相对路径。通过 SMB 导出的文件系统仅使用文件系统的资源名称，但每个文件系统仍需要在系统上的某个位置有一个唯一挂载点。
GUID	LUN 的 SCSI GUID。有关详细信息，请参见 ""Shares" (共享资源) > "Shares" (共享资源) > "Protocols" (协议) - BUI 页面 [290]。

以下工具可用于每个共享资源：

表 12-3 BUI "Shares" (共享资源) > "Shares" (共享资源) 图标

图标	说明
	将某个共享资源移动到其他项目。如果项目面板没有展开，这将会自动展开面板，直到该共享资源落入某个项目中。
	编辑单个共享资源（也可通过双击行来编辑）。
	销毁共享资源。系统将会提示您确认此项操作，因为这将会销毁共享资源中的所有数据，并且无法撤消。

编辑共享资源

要编辑某个共享资源，请单击铅笔图标或双击共享资源列表中的行。此操作将选择该共享资源，并显示多个不同的选项卡供选择来编辑共享资源的属性。在每个选项卡区域中可以找到完整的功能集合：

- [""General" \(常规\) "](#) [284]
- [""Protocols" \(协议\) "](#) [290]
- [""Access" \(访问\) "](#) [297]
- [""Snapshots" \(快照\) "](#) [303]
- [第 13 章 复制](#)

共享资源的名称显示在项目面板右侧的左上角。名称的第一个组成部分是包含的项目，单击项目名称将导航到[[Shares:Projects]项目详细信息]]。也可以通过单击共享资源名称并在输入框中输入新文本来更改共享资源的名称。系统将要求您确认该项操作，因为这将需要断开共享资源的活动客户机的连接。

使用情况统计信息

视图的左侧（展开时位于项目面板下面）是一个表，用于说明当前空间使用情况统计信息。这些统计信息针对特定共享资源（编辑某个共享资源时）或针对整个池（查看共

享资源列表时)。如果任何属性为零，则表中将排除该属性。显示下列使用情况统计信息：

- Available space (可用空间) – 此统计信息在标题中隐式显示为以容量百分比表示的容量。可用空间反映共享资源或项目的任何配额或池的绝对容量。此处显示的数字是使用的总空间和可用空间量之和。
- Referenced Data (引用数据) – 数据引用的数据量。除了必不可少的元数据外，还包括所有文件系统数据或 LUN 块。通过压缩，该值可能远小于共享资源中包含的数据的逻辑大小。如果共享资源是快照的克隆，则该值可能小于理论上它可包含的物理存储，并且可能为零。
- Snapshot Data (快照数据) – 共享资源的所有快照 (包括任何项目快照) 使用的空间量。此大小不等于所有快照占用的唯一空间之和。由多个快照引用的块不包括在每个快照的使用量统计信息中，但将显示在共享资源的快照数据总计中。
- Unused Reservation (未使用的预留空间) – 如果某个文件系统设置了预留空间，则该值表示该文件系统预留的剩余空间量。没有为 LUN 设置该值。ZFSSA 会阻止其他共享资源占用此空间，从而确保该文件系统有足够的空间。如果预留空间不包括快照，则在创建快照时必须有足够的空间才能覆盖整个快照。有关预留空间的更多信息，请参见“常规属性” [284] 一节。
- Total Space (总空间) – 引用数据、快照数据和未使用的预留空间之和。

静态属性

在编辑特定的共享资源时，共享资源视图的左侧还将显示静态 (创建时) 属性。这些属性是在创建时设置的，并且一旦设置便无法修改。显示下列静态属性：

- Compression Ratio (压缩率) – 如果启用了压缩，这将显示共享资源当前达到的压缩率。它表示为一个倍数。例如，2x 的压缩率表示此数据占用的空间是未压缩内容的一半。有关压缩和可用算法的更多信息，请参见“常规属性” [284] 一节。
- Case Sensitivity (区分大小写) – 控制目录查找是否区分大小写。它支持以下选项：

BUI 值	CLI 值	说明
Mixed (混合)	mixed (混合)	是否区分大小写取决于所使用的协议。对于 NFS、FTP 和 HTTP，查找区分大小写。对于 SMB，查找不区分大小写。这是默认设置，优先遵循各个协议的规定而不是在协议间保持一致性。使用此模式时，可能会出现以下情况：创建了多个文件，通过区分大小写的协议访问时这些文件不同，但通过 SMB 访问时发生冲突。在这种情况下，SMB 服务器将创建冲突项的“改编”版本来唯一地标识文件名。

BUI 值	CLI 值	说明
Insensitive (不区分大小写)	insensitive (不区分大小写)	所有查找都不区分大小写, 即使是使用传统上区分大小写的协议 (如 NFS)。对这些协议的客户端而言, 这会导致混淆, 但可以防止客户端创建冲突的名称, 名称冲突会导致通过 SMB 进行访问时使用改编的名称。此设置只能在下列情况下使用: SMB 是主协议, 将备用协议视为次要协议, 而且对预期标准的遵循不是问题。
Sensitive (区分大小写)	sensitive (区分大小写)	所有查找都区分大小写, 即使是使用传统上不区分大小写的 SMB。通常不应使用此设置, 因为 SMB 服务器会通过改编的名称来处理名称冲突, 并可能导致 Windows 应用程序行为异常。

- **Reject non UTF-8 (拒绝非 UTF-8)** – 此设置强制为所有文件和目录实施 UTF-8 编码。设置后, 尝试使用无效的 UTF-8 编码来创建文件或目录将失败。这只会影响 NFSv3, 该协议中编码不是按标准定义的。NFSv4 始终使用 UTF-8 和 SMB 来协商相应的编码。此设置通常应为 "on", 否则 SMB (SMB 必须知道编码及其他信息才能执行区分大小写的对比) 无法对使用无效 UTF-8 编码创建的文件名进行解码。只有在以下情况下才能将其设置为 "off": 在已存在的 NFSv3 部署中, 将客户端配置为使用不同的编码。如果 NFSv3 客户端创建了不是有效的 UTF-8 编码的文件或目录, 则在此属性设置为 "off" 的情况下启用 SMB 或 NFSv4 会产生不确定的结果。如果 "normalization" 属性设置为 "none" 以外的其他任何值, 则此属性必须设置为 "on"。
- **Normalization (标准化)** – 此设置控制对文件系统和目录执行哪种 Unicode 标准化 (如果有)。Unicode 支持用不同的编码表示相同的逻辑名称。如果没有标准化, 则所存储的磁盘上名称将不同, 根据创建文件以及访问文件的方式, 使用备用形式之一的查找将失败。如果此属性设置为 "none" (默认值) 以外的其他任何值, "Reject non UTF-8" (拒绝非 UTF-8) 属性必须也设置为 "on"。有关标准化执行方式以及各种形式执行方式的更多信息, 请参见有关 Unicode 标准化的 Wikipedia 条目。

BUI 值	CLI 值	说明
None	none	不执行标准化。
Form C	formC	标准形式规范化组合 (<i>Normalization Form Canonical Composition, NFC</i>) – 字符按规范化等值分解然后重组。
Form D	formD	标准形式规范化分解 (<i>Normalization Form Canonical Decomposition, NFD</i>) – 字符按规范化等值分解。

BUI 值	CLI 值	说明
Form KC	formKC	标准形式兼容性组合 (Normalization Form Compatibility Composition, NFKC) – 字符按兼容性等值分 解，然后按规范化等值重组。
Form KD	formKD	标准形式兼容性分解 (Normalization Form Compatibility Decomposition, NFKD) – 字符按兼容性等值分 解。

- Volume Block Size (卷块大小) – LUN 的原始块大小。可以是 512 字节到 1M 之间的 2 的任何次幂，默认值为 8K。
- Origin (来源) – 如果这是一个克隆，则这是克隆所依据的快照的名称。
- Data Migration Source (数据迁移源) – 设置后，该文件系统会在本地或通过 NFS 主动对现有的文件系统执行影子迁移操作。有关数据迁移的详细信息，请参见有关第 14 章 影子迁移一节。

共享资源项目面板

在 BUI 中，始终通过视图左侧的项目面板来显示可用的一组项目。要展开或折叠项目面板，请单击 "Projects" (项目) 标题栏中的三角形。

表 12-4 项目面板图标

图标	说明
	展开项目面板
	折叠项目面板

从面板中选择一个项目将导航至选定项目的“项目” [312]视图。在共享资源列表内的某个行上单击移动工具时，此项目面板也会自动展开。然后您可以拖放共享资源以将其在项目之间移动。项目面板还提供了快捷方式，用于创建新项目以及返回到所有项目中的共享资源列表。单击 "All" (全部) 文本等效于在导航栏中选择 "Shares" (共享资源) 项。

项目面板为项目数量相对较少的系统提供了方便。该面板的设计目的不在于充当可管理大量项目的主界面。对于此任务，请参见“项目” [312]视图。

▼ 创建共享资源

1. 要查看某个项目或所有项目中的共享资源，请转至 "Shares" (共享资源) > "Shares" (共享资源)。
2. 选择 "Filesystems" (文件系统) 或 "LUNs" (LUN)。
3. 单击 "Filesystems" (文件系统) 或 "LUNs" (LUN) 旁边的加号图标。
此时会显示 "Create Filesystem" (创建文件系统) 或 "Create LUN" (创建 LUN) 对话框。
4. 在 "Create Filesystem" (创建文件系统) 或 "Create LUN" (创建 LUN) 对话框中，选择或键入您要使用的属性。

每种类型的共享资源的属性在以下位置进行定义：

对于文件系统：

- “用户” [297]
- “组” [297]
- “权限” [297]
- “挂载点” [284]
- 第 12 章 共享资源、项目和模式 (仅创建时)
- 第 12 章 共享资源、项目和模式 (仅创建时)
- 第 12 章 共享资源、项目和模式 (仅创建时)

对于 LUN：

- “卷大小” [284]
- “Thin provisioned (瘦置备)” [284]
- 第 12 章 共享资源、项目和模式 (仅创建时)

在 CLI 中使用 "Shares" (共享资源) > "Shares" (共享资源)

共享资源 CLI 位于 shares 下

导航

在选择共享资源之前，必须首先选择一个项目（包括默认项目）：

```
clownfish:> shares
```

```

clownfish:shares> select default
clownfish:shares default> select foo
clownfish:shares default/foo> get
Properties:
    aclinherit = restricted (inherited)
    aclmode = discard (inherited)
    atime = true (inherited)
    casesensitivity = mixed
    checksum = fletcher4 (inherited)
    compression = off (inherited)
    compressratio = 100
    copies = 1 (inherited)
    creation = Mon Oct 13 2009 05:21:33 GMT+0000 (UTC)
    mountpoint = /export/foo (inherited)
    normalization = none
    quota = 0
    quota_snap = true
    readonly = false (inherited)
    recordsize = 128K (inherited)
    reservation = 0
    reservation_snap = true
    secondarycache = all (inherited)
    nbmand = false (inherited)
    sharesmb = off (inherited)
    sharenfs = on (inherited)
    snapdir = hidden (inherited)
    snaplabel = project1:share1
    utf8only = true
    vscan = false (inherited)
    sharedav = off (inherited)
    shareftp = off (inherited)
    space_data = 43.9K
    space_unused_res = 0
    space_snapshots = 0
    space_available = 12.0T
    space_total = 43.9K
    root_group = other
    root_permissions = 700
    root_user = nobody

```

共享资源操作

可以通过选择项目并执行 `filesystem` 或 `lun` 命令来创建共享资源。在提交更改之前，可以根据需要对属性进行修改：

```

clownfish:shares default> filesystem foo
clownfish:shares default/foo (uncommitted)> get
    aclinherit = restricted (inherited)
    aclmode = discard (inherited)
    atime = true (inherited)
    checksum = fletcher4 (inherited)
    compression = off (inherited)
    copies = 1 (inherited)
    mountpoint = /export/foo (inherited)

```

```

        quota = 0 (inherited)
        readonly = false (inherited)
        recordsize = 128K (inherited)
        reservation = 0 (inherited)
        secondarycache = all (inherited)
        nbmand = false (inherited)
        sharesmb = off (inherited)
        sharenfs = on (inherited)
        snapdir = hidden (inherited)
        snaplabel = project1:share1
        vscan = false (inherited)
        sharedav = off (inherited)
        shareftp = off (inherited)
        root_group = other (default)
        root_permissions = 700 (default)
        root_user = nobody (default)
        casesensitivity = (default)
        normalization = (default)
        utf8only = (default)
        quota_snap = (default)
        reservation_snap = (default)
        custom:int = (default)
        custom:string = (default)
        custom:email = (default)
clownfish:shares default/foo (uncommitted)> set sharenfs=off
        sharenfs = off (uncommitted)
clownfish:shares default/foo (uncommitted)> commit
clownfish:shares default>

```

可以在共享资源上下文中使用 `destroy` 命令来销毁共享资源：

```

clownfish:shares default/foo> destroy
This will destroy all data in "foo"! Are you sure? (Y/N)
clownfish:shares default>

```

可以在项目上下文中使用 `rename` 命令来重命名共享资源：

```

clownfish:shares default> rename foo bar
clownfish:shares default>

```

可以在项目上下文中使用 `move` 命令在项目之间移动共享资源：

```

clownfish:shares default> move foo home
clownfish:shares default>

```

在选择了特定项目或共享资源之后，可以通过 `users` 或 `groups` 命令来管理用户和组使用情况和配额。有关如何管理用户和组配额的更多信息，请参见[“空间管理” \[265\]](#)一节。

"Shares" (共享资源) > "Shares" (共享资源) CLI 属性

下面显示了在 CLI 中提供的属性及其在 BUI 中对应的属性。可以使用标准 CLI 命令 `get` 和 `set` 来设置这些属性。此外，还可以使用 `unset` 命令从父项目中继承这些属性。

表 12-5 "Shares" (共享资源) > "Shares" (共享资源) CLI 属性

CLI 名称	"类型" [262]	BUI 名称	BUI 位置
aclinherit	继承	"ACL inheritance behavior (ACL 继承行为)" [297]	Access (访问)
aclmode	继承	"ACL behavior on mode change (模式更改时的 ACL 行为)" [297]	Access (访问)
atime	继承	"Update access time on read (读取时更新访问时间)" [284]	General (常规)
casesensitivity	创建时	第 12 章 共享资源、项目和模式	Static (静态)
checksum	继承	第 12 章 共享资源、项目和模式	General (常规)
compression	继承	"Data compression (数据压缩)" [284]	General (常规)
compresratio	只读	第 12 章 共享资源、项目和模式	Static (静态)
copies	继承	"Additional replication (额外复制)" [284]	General (常规)
creation	只读	-	-
dedup	继承	"Data deduplication (重复数据删除)" [284]	General (常规)
exported	继承, 仅限复制包	第 13 章 复制	General (常规)
fixednumber	LUN 本地	"Initiator group (启动器组)" [290]	Protocols (协议)
initiatorgroup	LUN 本地	"Initiator group (启动器组)" [290]	Protocols (协议)
logbias	继承	"Synchronous write bias (同步写入偏向)" [284]	General (常规)
lunumber	LUN 本地	"LU number (LU 编号)" [290]	Protocols (协议)
lunguid	只读, LUN 本地	"GUID" [290]	Protocols (协议)
mountpoint	继承	"Mountpoint (挂载点)" [284]	General (常规)
nbmand	继承	"Non-blocking mandatory locking (非	General (常规)

CLI 名称	“类型” [262]	BUI 名称	BUI 位置
		阻塞强制性锁定) ” [284]	
nodestroy	继承	“Prevent destruction (阻止销毁) ” [284]	General (常规)
normalization	创建时	第 12 章 共享资源、项目和模式	Static (静态)
origin	只读	第 12 章 共享资源、项目和模式	Static (静态)
quota	空间管理	“quota (配额) ” [265]	General (常规)
quota_snap	空间管理	“Quota / Include snapshots (配额/包含快照) ” [265]	General (常规)
readonly	继承	“Read-only (只读) ” [284]	General (常规)
recordsize	继承	“Database record size (数据库记录大小) ” [284]	General (常规)
reservation	空间管理	“Reservation (预留空间) ” [265]	General (常规)
reservation_snap	空间管理	“Reservation / Include snapshots (预留空间/包含快照) ” [265]	General (常规)
root_group	文件系统本地	“Group (组) ” [297]	Access (访问)
root_permissions	文件系统本地	“Permission (权限) ” [297]	Access (访问)
root_user	文件系统本地	“User (用户) ” [297]	Access (访问)
rstchown	继承	“Restrict ownership change (限制所有权更改) ” [284]	General (常规)
secondary cache	继承	“Cache device usage (高速缓存设备使用) ” [284]	General (常规)
shadow	创建时	第 14 章 影子迁移	Static (静态)
sharedav	继承	“Protocols / HTTP / Share mdoe (协议/HTTP/共享模式) ” [290]	Protocols (协议)
shareftp	继承	“Protocols / FTP / Share mode (协议/FTP/共享模式) ” [290]	Protocols (协议)

CLI 名称	"类型" [262]	BUI 名称	BUI 位置
share nfs	继承	"Protocols / NFS / Share mode (协议/NFS/共享模式) " [290]	Protocols (协议)
share smb	继承	"Protocols / SMB / Resource name (协议/SMB/资源名称) " [290]	Protocols (协议)
snapdir	继承	".zfs/snapshot visibility (.zfs/snapshot 可见性) " [303]	Snapshots (快照)
snaplabel	继承	"Scheduled snapshot label (调度的快照标签) " [303]	Snapshots (快照)
space_available	只读	第 12 章 共享资源、项目和模式	使用情况
space_data	只读	第 12 章 共享资源、项目和模式	使用情况
space_snapshots	只读	第 12 章 共享资源、项目和模式	使用情况
space_total	只读	第 12 章 共享资源、项目和模式	使用情况
space_unused_res	只读	第 12 章 共享资源、项目和模式	使用情况
sparse	LUN 本地	"Thin provisioned (瘦置备) " [284]	General (常规)
targetgroup	LUN 本地	"Target group (目标组) " [290]	Protocols (协议)
utf8only	创建时	"Reject non UTF-8 (拒绝非 UTF-8) "	Static (静态)
volblocksize	创建时	第 12 章 共享资源、项目和模式	Static (静态)
vscan	继承	"Virus scan (病毒扫描) " [284]	General (常规)

"Shares" (共享资源) > "Shares" (共享资源) > "General" (常规) – BUI 页面

在 BUI 中，这部分内容用于控制共享资源的整体设置，这些设置与任何特定协议无关，也与访问控制或快照无关。虽然 CLI 将所有属性组合到了单个列表中，但本节介绍了在这两种上下文中属性的行为。

这些是可以从项目中继承或显式在共享资源上设置的标准属性。BUI 只允许一次性继承所有属性，但 CLI 允许分别继承各个属性。

有关如何将这些属性映射到 CLI 的信息，请参见[“在 CLI 中使用 “Shares” \(共享资源\) > “Shares” \(共享资源\)” \[278\]](#)。

Space Usage (空间使用情况)

存储池中的空间供所有共享资源使用。尽管可以以共享资源为单位实施空间限制，但文件系统可以根据需要动态地扩展或收缩。可以以文件系统为单位实施配额和预留空间。还可以以用户和组为单位实施配额。有关管理文件系统的空间使用情况（包括配额和预留空间）的详细信息，请参见[“空间管理” \[265\]](#)一节。

Volume Size (卷大小)

通过 iSCSI 进行导出时 LUN 的逻辑大小。此属性只对 LUN 有效。此属性控制 LUN 的大小。默认情况下，LUN 会预留足够的空间来完全填充该卷。有关更多信息，请参见[“瘦置备” \[284\]](#)属性。将处于活动状态的 LUN 导入到客户机时，如果更改其大小，可能会产生不确定的结果。可能需要客户机重新连接并/或导致基于 LUN 的文件系统中发生数据损坏。在尝试此操作之前，请先确定适合您的特定 iSCSI 客户机的最佳做法。

Thin Provisioned (瘦置备)

控制是否为卷预留空间。此属性只对 LUN 有效。默认情况下，LUN 会预留恰好足以完全填充卷的空间。这将确保客户机不会不合时宜地出现空间不足的错误。此属性使得卷的大小超过可用空间的大小。设置后，LUN 将只会占用已写入到 LUN 的空间。尽管这允许对 LUN 进行瘦置备，但大多数文件系统不希望底层设备出现“空间不足”的状况，并且一旦共享资源空间不足，则可能会导致在客户机上发生不稳定和/或数据损坏的情况。

当未设置此属性时，卷大小的行为与排除了快照的预留空间相同。因此，它具有相同的症状，包括当快照在理论上可以分化到超过可用空间量的点时创建快照会失败。有关更多信息，请参阅[“项目 – 预留空间” \[318\]](#)。

Mountpoint (挂载点)

挂载文件系统的位置。此属性只对文件系统有效。

对于挂载点属性，有以下限制：

- 必须位于 `/export` 下。
- 不得与其他共享资源冲突。
- 不得与群集对等节点上的其他共享资源冲突，以便能够正确地进行故障转移。

继承挂载点属性时，会将当前数据集名称附加到项目的挂载点设置，并以斜线 (`/`) 连接。例如，如果 `"home"` 项目的挂载点设置是 `/export/home`，则 `"home/bob"` 将继承挂载点 `/export/home/bob`。

SMB 共享资源是通过其资源名称导出的，并且无法通过协议查看挂载点。不过，即使是仅 SMB 共享资源也必须在 ZFSSA 上拥有有效的唯一挂载点。

尽管有一些限制，但可以在其他共享资源下嵌入挂载点。有关详细信息，请参见[“文件系统名称空间” \[271\]](#)一节。

Read only (只读)

控制文件系统内容是否为只读内容。此属性只对文件系统有效。无论协议设置如何，只读文件系统的内容都无法修改。该设置不会影响重命名、销毁或更改文件系统属性的能力。此外，当文件系统为只读系统时，无法更改[“访问控制” \[297\]](#)属性，因为它们需要修改文件系统根目录的属性。

Update access time on read (读取时更新访问时间)

控制在读取时是否更新文件的访问时间。此属性只对文件系统有效。POSIX 标准要求文件的访问时间要正确地反映最后一次读取文件的时间。这需要执行到底层文件系统的写入，即使对主要为只读内容的工作负荷也是如此。对于主要由对大量文件的读取组成的工作集，禁用此属性可以改进性能，但代价是牺牲了标准符合性。这些更新将以异步方式执行并组合在一起，因此其影响应当表现不出来，除非在负荷过重的情况下。

Non-blocking mandatory locking (非阻塞强制性锁定)

控制是否通过 POSIX 语义实施 SMB 锁定语义。此属性只对文件系统有效。默认情况下，文件系统会根据 POSIX 标准实施文件行为。这些标准从根本上与 SMB 协议所需的行为不兼容。对于主协议为 SMB 的共享资源，应始终启用此选项。更改此属性会要求所有客户机先断开连接，然后重新连接。

Data deduplication (重复数据删除)

控制是否清除数据的重复副本。重复数据删除是基于块的，在池范围内同步进行，并且可以以项目或共享资源为单位启用。通过在项目或共享资源的常规属性屏幕上选中 "Data Deduplication" (重复数据删除) 复选框可以启用该功能。重复数据删除率将显示在状态显示板中的使用量区域中。

在启用重复数据删除情况下写入的数据将被输入到以数据校验和为索引的重复数据删除表中。重复数据删除强制使用加密性强的 SHA-256 校验和。随后的写入将识别重复数据，并且只在磁盘上保留现有副本。只能在大小相同的块间对以相同的记录大小写入的数据执行重复数据删除。与往常一样，为获得最佳结果，请将记录大小设置为使用该数据的应用程序的记录大小；对于流化处理工作负荷，请使用较大的记录大小。

如果您的数据未包含任何重复项，则启用 "Data Deduplication" (重复数据删除) 将会增加开销 (更加密集占用 CPU 的校验和以及磁盘上重复数据删除表条目)，不会带来任何益处。如果您的数据确实包含重复项，则无论给定的块出现多少次，启用 "Data Deduplication" (重复数据删除) 后都将只存储一个副本，从而可以节省空间。重复数据删除必然会影​​响性能，以至于校验和的计算更加昂贵，而且必须访问和维护重复数据删除表的元数据。

请注意，重复数据删除对所计算的共享资源大小不会产生任何影响，但会影响用于池的空间量。例如，如果两个共享资源包含相同的 1GB 文件，则每个共享资源的大小将显示为 1GB，但是池的总大小也正好是 1GB，而且重复数据删除率将报告为 2x。

性能警告：重复数据删除要求在写入块或释放块时修改重复数据删除表 (由其本质所决定)。如果重复数据删除表无法完全装入 DRAM 中，则写入和释放可能会导致发生先前从未出现过的大量随机读取活动。因此，启用重复数据删除功能会对性能产生严重影响。而且，在某些情况下 (特别是在删除共享资源或快照的情况下)，在整个池范围内可能都会感觉到启用重复数据删除所导致的性能下降。一般情况下，除非已知某个共享资源的重复数据比率很高而且重复的数据和引用该数据的表可以很合身地装入 DRAM 中，否则不建议启用重复数据删除。要确定重复数据删除是否会对性能造成负面影响，请启用第 8 章 设置 ZFSSA 首选项，使用《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的 "Analytics" 来测量 "ZFS DMU operations broken down by DMU object type" (按 DMU 对象类型细分的 ZFS DMU 操作数)，然后检查与 ZFS 操作相比是否存在比率较高的持续 DDT (Data Duplication Table, 数据重复表) 操作。如果出现此情况，则表明更多的是 I/O 而不是文件 I/O 在为重复数据删除表提供服务。

Data compression (数据压缩)

控制在将数据写入磁盘之前是否对其进行压缩。共享资源可以先压缩数据，然后再将其写入存储池。这可以实现较高的存储利用率，但会增加 CPU 的利用率。默认情况下，不会进行压缩。如果压缩不能获得最低限度的空间节省，则不会将其提交到磁盘，以避免在读回数据时进行不必要的解压缩。在选择压缩算法之前，建议您先执行一些必要的性能测试并测量所达到的压缩比率。

BUI 值	CLI 值	说明
Off (关闭)	关	不执行压缩
LZJB (最快)	lzjb	一种简单的运行长度编码，仅适用于十分简单的输入，但不会占用很多 CPU。
GZIP-2 (快速)	gzip-2	gzip 压缩算法的轻量级版本。
GZIP (默认)	gzip	标准 gzip 压缩算法。
GZIP-9 (最佳压缩)	gzip-9	使用 gzip 可达到的最大压缩程度。这将会占用大量 CPU，而且经常只能勉强达到目标。

Checksum (校验和)

控制用于数据块的校验和。在 ZFSSA 上，将在磁盘上对所有数据执行校验和运算，这样可以避免传统的缺陷（尤其是虚读和虚写）。这使得系统可以检测到设备返回的无效数据。对于正常操作，默认的校验和 (fletcher4) 就足够了，但偏执的用户可能会增大校验和的长度，这会额外增加 CPU 负荷。由于始终使用相同的算法对元数据执行校验和运算，因此这只影响用户数据（文件或 LUN 块）。

BUI 值	CLI 值	说明
Fletcher 2 (传统)	fletcher2	16 位 fletcher 校验和
Fletcher 4 (标准)	fletcher4	32 位 fletcher 校验和
SHA-256 (特强)	sha256	SHA-256 校验和

Cache Device Usage (高速缓存设备使用)

控制是否将高速缓存设备用于共享资源。默认情况下，所有数据集都将使用系统上的任何高速缓存设备。高速缓存设备是作为存储池的一部分配置的，它提供了一个额外的高速缓存层，以实现更快的分层访问。有关高速缓存设备的详细信息，请参见[第 5 章 存储配置](#)一节。此属性与存储池中当前是否配置了高速缓存设备无关。例如，即使在没有高速缓存设备的情况下，该属性仍可以设置为 "all" (全部)。如果以后添加了任何此类设备，共享资源将自动利用附加的性能。此属性不会影响主 (DRAM) 高速缓存的使用。

BUI 值	CLI 值	说明
All data and metadata (所有数据和元数据)	all	缓存所有普通文件或 LUN 数据以及所有元数据。

BUI 值	CLI 值	说明
Metadata only (仅元数据)	metadata	仅在高速缓存设备上保留元数据。这可以快速遍历目录结构，但检索文件内容可能需要从数据设备读取数据。
Do not use cache devices (不使用高速缓存设备)	none	不在高速缓存设备上缓存此共享源中的数据。数据只缓存在主高速缓存中或存储在数据设备上。

Synchronous Write Bias (同步写入偏向)

此设置控制在为同步写入提供服务时的行为。默认情况下，系统将针对延迟优化同步写入，利用日志设备提供快速响应时间。在具有多个不相交文件系统的系统中，这可能会导致对日志设备的争用，从而增加了所有使用者的延迟。即使在多个文件系统要求同步语义的情况下，也可能会出现某些文件系统比其他文件系统对延迟更敏感的情况。

具有单独日志的数据库就是一个常见的案例。日志对延迟极其敏感，尽管数据库本身也要求同步语义，但它是负载较重的带宽，对延迟不敏感。在此环境中，在主数据库中将此属性设置为 "throughput" (吞吐量) 而将日志文件系统保留为 "latency" (延迟) 会使性能显著改进。即使没有日志设备时，此设置也将改变系统的行为，虽然影响不是很明显。

"Synchronous write bias" (同步写入偏向) 设置可以由 Oracle Intelligent Storage Protocol 绕过。Oracle Intelligent Storage Protocol 可以使用由 Oracle Database NFSv4 客户机提供的写入偏向值，而不使用在文件系统中定义的写入偏向。由 Oracle Database NFSv4 客户机发送的写入偏向值仅用于该写入请求。有关更多信息，请参阅["Oracle Intelligent Storage Protocol" \[428\]](#)。

BUI 值	CLI 值	说明
Latency (延迟)	latency	如果有专用的日志设备，则利用该设备针对延迟对同步写入进行优化。
Throughput (吞吐量)	throughput	针对吞吐量对同步写入进行优化。数据将写入到主数据磁盘而不是日志设备，并且写入将以针对系统的总带宽进行了优化的方式执行。

Database Record Size (数据库记录大小)

为文件系统中的文件指定建议的块大小。该属性仅对文件系统有效，设计用于对大小固定的记录中的文件进行访问的数据库工作负荷。系统会根据为典型的访问模式优化的内部算法来自动调整块大小。

对于创建很大的文件但访问较小的随机块中的文件的数据库而言，这些算法可能不是最优的。将记录大小指定为大于或等于数据库的记录大小可以显著提高性能。强烈建议不要将该属性用于一般用途的文件系统，否则可能会对性能产生不利影响。

默认记录大小为 128 KB。指定的大小必须是 2 的若干次幂，并且必须大于或等于 512 字节同时小于或等于 1 MB。更改文件系统的记录大小只影响以后创建的文件，现有文件和接收的数据不受影响。

注意：如果大于 128K 的块大小用于项目或共享资源，则将这些项目或共享资源复制到不支持大型块的系统将失败。

"Database record size" (数据库记录大小) 设置可以由 Oracle Intelligent Storage Protocol 绕过。Oracle Intelligent Storage Protocol 可以使用由 Oracle Database NFSv4 客户机提供的块大小值，而不使用在文件系统中定义的记录大小。由 Oracle Database NFSv4 客户机提供的块大小只能在创建新数据库文件或表时应用。现有文件和表的块大小将不进行更改。有关更多信息，请参阅“[Oracle Intelligent Storage Protocol](#)” [428]。

Additional replication (额外复制)

控制除存储池的任何冗余之外为每个块存储的副本数。元数据始终随多个副本一同存储，但是此属性允许向数据块应用相同的行为。存储池将尽量在不同的设备上存储这些额外的块，但无法确保会做到这一点。此外，如果丢失了一个完整的逻辑设备（RAID 条带、镜像对等等），则将无法导入存储池。此属性不能替代存储池中的正确复制，但可以令偏执的管理员放心。

BUI 值	CLI 值	说明
Normal (Single Copy) (常规 (单个副本))	1	默认行为。存储数据块的单个副本。
Two Copies (两个副本)	2	存储每个数据块的两个副本。
Three Copies (三个副本)	3	存储每个数据块的三个副本。

Virus scan (病毒扫描)

控制是否对此文件系统进行病毒扫描。此属性只对文件系统有效。此属性设置与病毒扫描服务的状态无关。即使已启用了病毒扫描服务，也必须使用此属性显式启用文件系统扫描。同样，即使该服务本身已关闭，仍可以为特定的共享资源启用病毒扫描。有关配置病毒扫描的更多信息，请参见“[病毒扫描](#)” [220] 一节。

Prevent Destruction (防止销毁)

当设置了此属性时，无法销毁共享资源或项目。这包括通过从属克隆销毁共享资源，销毁项目内的共享资源或销毁复制数据包。不过，这不会影响通过复制更新销毁的共享资源。如果在作为复制源的 ZFSSA 上销毁了某个共享资源，即使设置了此属性，目标设备上的对应共享资源也将被销毁。

要销毁共享资源，首先必须要做的是显式禁用此属性。默认情况下，此属性处于禁用状态。

Restrict Ownership Change (限制所有权更改)

默认情况下，除了 root 用户（在具有启用了 root 导出的合适客户机上）之外，其他人无法更改文件的所有权。可以通过关闭此属性来以文件系统或项目为单位禁用此属性。禁用后，文件或目录的所有者可以更改文件的所有权，实际上是允许用户“放弃”自己的文件。更改所有权时，将剥离任何 setuid 或 setgid 位，以防止用户通过此操作提升特权。

定制属性

可以根据需要添加定制属性，以便将用户定义的标记附加到项目和共享资源。有关更多信息，请参见[“模式” \[322\]](#)。

"Shares" (共享资源) > "Shares" (共享资源) > "Protocols" (协议) – BUI 页面

共享资源协议

每个共享资源都有特定于协议的属性，这些属性定义了不同的协议针对该共享资源的行为。可以为每个共享资源定义这些属性，也可以从共享资源的项目继承这些属性。“NFS” [185]、“SMB” [191]、“HTTP” [208]和“FTP” [206]属性仅应用于文件系统，而“iSCSI” [189]属性仅应用于 LUN。

在 BUI 中，每个协议都显示使用该协议的客户机将用来引用共享资源的路径。例如，服务器 "twofish" 上的文件系统 "fs0" 将通过以下位置提供：

表 12-6 共享资源协议

协议	所在位置
NFS	twofish:/export/fs0
SMB	\\twofish\fs0
HTTP	//twofish/shares/export/fs0/
FTP	ftp://twofish/export/fs0/
SFTP	/export/fs0/

对于 iSCSI，启动器可以通过[第 6 章 存储区域网络配置](#)中所述的机制之一来发现目标。

共享资源协议 – NFS

表 12-7 共享资源协议 – NFS 属性

BUI 属性	CLI 属性	说明
Share mode (共享模式)	off/ro/rw	确定共享资源是只可用于读取、可用于读写，还是不可用于两者中的任何一个。在 CLI 中，"on" 是 "rw" 的别名。
Disable setuid/setgid file creation (禁止创建 setuid/setgid 文件)	nosuid	如果选择了此选项，则客户机将无法在设置了 setuid (S_ISUID) 和 setgid (S_ISGID) 位的情况下创建文件，也无法通过 chmod(2) 系统调用在现有文件上启用这些位。
Prevent clients from mounting subdirectories (阻止客户机挂载子目录)	nosub	如果选择了此选项，将会阻止客户机直接挂载子目录。将强制它们挂载共享资源的根目录。注意：这只适用于 NFSv2 和 NFSv3 协议，不适用于 NFSv4。
Anonymous user mapping (匿名用户映射)	anon	除非为特定客户机启用了 "root" 选项，否则，客户机上的 root 用户将被视为未知用户，且该用户对共享资源的文件的所有访问尝试将被视为具有此 uid 的用户的尝试。然后，将正常评估文件的访问位和 ACL。
Character encoding (字符编码)	请参见下文	为所有客户机设置默认的字符集。有关更多信息，请参见有关字符集编码的章节。
Security mode (安全模式)	请参见下文	为所有客户机设置安全模式。

可以为客户端或客户端集合定义整体共享模式的例外。当某个客户端尝试访问时，将根据列表中与该客户端匹配的第一个例外向其授予访问权限；或者，如果不存在这样的例外，则将根据上面定义的全局共享模式向其授予访问权限。可以使用以下三种类型之一定义这些客户端集合：

表 12-8 客户端集合类型

类型	CLI 前缀	说明	示例
主机 (FQDN) 或网络组	无	其 IP 地址解析为指定的全限定名称的单个客户端，或包含客户端的 IP 地址解析为的全限定名称的网络组	caji.sf.example.com
DNS 域	.	其 IP 地址解析为以此后缀结尾的全限定名称的所有客户端	sf.example.com
网络	@	其 IP 地址位于指定的 IP 子网（以 CIDR 表示法表示）内的所有客户端	192.168.20.0/22

对于每个指定的客户端或客户端集合，您将表达两个参数：允许客户端对共享资源具有只读访问权限还是具有读写访问权限，以及是将客户端上的 root 用户视为 root 用户（如果被选中）还是视为未知用户。

如果使用网络组，则将首先通过“NIS” [223]（如果已启用），然后再通过“LDAP” [225]（如果已启用）解析这些网络组。如果使用 LDAP，则必须在默认位置 ou=Netgroup,(Base DN) 查找网络组，并且必须使用标准模式。网络组条目的用户名组件通常不会影响 NFS；只有主机名有意义。网络组中包含的主机名必须是规范化的，如果使用 DNS 进行解析，则必须是全限定的。也就是说，NFS 子系统将尝试验证请求方客户端的 IP 地址是否可以解析为与指定的 FQDN 或与指定的网络组之一的某个成员匹配的规范化主机名。该匹配必须为严格匹配，包括所有域组件；否则，将不会匹配该例外并将尝试下一个例外。有关主机名解析的更多信息，请参见“DNS” [239]。网络组的管理可能比较复杂；如有可能，请考虑换用 IP 子网规则或 DNS 域规则。

从 2013.1.0 软件发行版开始，Unix 客户端用户最多可以属于 1024 个组，而性能丝毫不减。在以前的发行版中，每个 Unix 客户端用户最多可以属于 16 个组。

共享资源协议 – CLI

在 CLI 中，所有 NFS 共享模式和例外都是通过针对 "sharenfs" 属性使用单个选项字符串指定的。该字符串是上表中值的逗号分隔的列表。作为针对 BUI 介绍的全局共享模式的对应体，该字符串应该以 "ro"、"rw" 或 "off" 之一开头。例如，

```
set sharenfs=ro
```

将所有客户机的共享模式设置为只读。所有客户机上的 root 用户将如同普通 "nobody" 用户一样访问共享资源上的文件。

此外，还将附加 "nosuid" 和 "anon" 中的其中一个选项或两者。请记住，在 CLI 中，必须将包含 "=" 字符的属性值加上引号。因此，要将所有未知用户的映射定义为 uid 153762，您可以指定

```
set sharenfs="ro,anon=153762"
```

可以通过附加 "option=collection" 格式的文本（其中 "option" 是 "ro"、"rw" 和 "root" 中的一个，定义将授予客户机集合的访问权限类型）来指定更多的例外。客户机集合是由上表中的前缀字符，以及一个 DNS 主机名/域名或 CIDR 网络号指定的。例如，要授予对 sf.example.com 域中的所有主机的读写权限，并授予对 192.168.44.0/24 网络中的主机的 root 访问权限，您可以使用

```
set sharenfs="ro,anon=153762,rw=.sf.example.com,root=@192.168.44.0/24"
```

在任何可以使用单个全限定主机名的位置中都可以使用网络组名称。例如，您可以通过以下方式允许对 "engineering" 网络组进行读写访问：

```
set sharenfs="ro,rw=engineering"
```

可以通过附加 "option=mode" 格式（其中，option 是 "sec"，mode 是 "sys"、"krb5"、"krb5:krb5i" 或 "krb5:krb5i:krb5p" 中的一个）的文本来指定安全模式。

```
set sharenfs="sec=krb5"
```

安全模式

安全模式是以共享资源为单位设置的并且可能会影响性能。下表介绍了 Kerberos 安全设置。

表 12-9 Kerberos 安全设置

设置	说明
krb5	通过 Kerberos V5 进行最终用户验证
krb5i	krb5 加完整性保护（数据包是防篡改的）
krb5p	krb5 加隐私保护（数据包是防篡改而经过加密的）

Kerberos 风格的组合可以在安全模式设置中指定。组合安全模式允许客户机使用所列出的任何 Kerberos 风格进行挂载。

表 12-10 安全模式设置

设置	菜单
sys	系统验证
krb5	仅限 Kerberos v5 – 客户机必须使用此风格进行挂载。
krb5:krb5i	Kerberos v5, 带有完整性 – 客户机可以使用所列出的任何风格进行挂载。
krb5i	仅限 Kerberos v5 完整性 – 客户机必须使用此风格进行挂载。
krb5:krb5i:krb5p	Kerberos v5, 带有完整性或隐私 – 客户机可以使用所列出的任何风格进行挂载。
krb5p	仅限 Kerberos v5 隐私 – 客户机必须使用此风格进行挂载。

有关使用 NFS 和 Kerberos 的更多信息，请参见：

- <http://www.ietf.org/rfc/rfc2623.txt> (<http://www.ietf.org/rfc/rfc2623.txt>) (NFSv2 和 NFSv3 安全性)
- <http://www.ietf.org/rfc/rfc3530.txt> (<http://www.ietf.org/rfc/rfc3530.txt>) (NFSv4 协议)

字符集编码

通常情况下，不指定用于文件名的字符集编码。NFSv3 和 NFSv2 协议没有指定字符集。NFSv4 应该使用 UTF-8，但不是所有的客户机都如此，而且服务器并未实施此限制。如果为共享资源禁用了 "UTF-8 only" 选项，则在不知道文件名编码的情况下会将文件名按原样写入文件系统。这意味着只有使用相同编码的客户机才能解释这些文件名。不过，SMB 要求将文件名存储为 UTF-8，以便在服务器端可以解释它们。这使得无法在允许通过 SMB 进行访问的同时支持任意客户机编码。

要支持这样的配置，可以在共享资源范围内或以客户机为单位设置字符集编码。支持以下字符集编码：

- cp932
- euc-cn
- euc-jp
- euc-jpms
- euc-kr
- euc-tw
- iso8859-1

- iso8859-2
- iso8859-5
- iso8859-6
- iso8859-7
- iso8859-8
- iso8859-9
- iso8859-13
- iso8859-15
- koi8-r
- shift_jis

默认行为是不指定字符集编码（直通）。BUI 允许通过标准例外列表机制选择字符集。在 CLI 中，每个字符集本身成为了具有一个或多个主机的选项，使用 "*" 表示共享资源范围内的设置。例如，以下示例：

```
set sharenfs="rw,euc-kr=*"
```

将以 "euc-kr" 作为默认编码来共享文件系统。请注意以下内容：

```
set sharenfs="rw,euc-kr=host1.domain.com,euc-jp=host2.domain.com"
```

为 "host1" 和 "host2" 之外的所有客户机使用默认编码，这两台客户机分别使用 "euc-kr" 和 "euc-jp"。主机列表的格式遵循其他 CLI NFS 选项的格式。

请注意，某些 NFS 客户机无法正确支持备用语言环境；有关详细信息，请查阅 NFS 客户机文档。

共享资源 – SMB

- Resource name (资源名称) – “SMB” [191] 客户机用来引用此共享资源的名称。资源名称 "off" 表示没有“SMB” [191] 客户机可以访问共享资源，资源名称 "on" 表示共享资源将随文件系统的名称一起导出。
- Enable Access-based Enumeration (启用基于访问的枚举) – 当启用了此选项时，将执行基于访问权限的枚举。基于访问权限的枚举根据客户机的凭证过滤目录条目。如果客户机无权访问某个文件或目录，则返回到客户机的条目列表中将省略该文件或目录。默认情况下不启用此选项。
- Is a DFS Namespace (是 DFS 名称空间) – 此属性指示是否将此共享资源置备为独立的“DFS 名称空间” [191]。
- Share-level ACL (共享资源级 ACL) – 此 ACL 与共享资源中文件或目录的 ACL 结合使用来确定该文件的有效权限。默认情况下，此 ACL 向每个人授予完全控制权限。此 ACL 在文件上的 ACL 之上又提供了一个访问控制层，以便实现更加复杂

的访问控制配置。只有在通过配置 SMB 资源名称导出文件系统后，才可设置此属性。如果未通过 SMB 协议导出文件系统，则设置共享资源级 ACL 不起任何作用。

同一系统上的两个“SMB” [191]共享资源不可具有相同的资源名称。从项目继承的资源名称具有特殊的行为，有关详细信息，请参见“项目” [312]一节。资源名称不得超过 80 个字符，可以包含除以下字符之外的任何字母数字字符：

```
" / \ [ ] : | < > + ; , ? * =
```

启用基于访问权限的枚举后，客户机可以看到它们无法打开的文件的目录条目。只有当客户机无权访问该文件时才会过滤掉目录条目。例如，如果客户机尝试打开某个文件进行读/写访问，但 ACL 只授予了读取访问权限，则打开请求将失败，不过条目列表中仍会包括该文件。

共享资源 – iSCSI

- Target group (目标组) – 通过其导出此 LUN 的目标。
- Initiator group(s) (启动器组) – 可以访问此 LUN 的启动器。从 2013.1.0 软件发行版开始，可以将多个启动器组分配给一个 LUN。编辑启动器组时，选中 "PERSIST" 复选框 (默认) 会保留相应启动器组的 LUN 编号。如果取消选中，则在 SAN 配置更改或重新引导之后，ZFSSA 可能重新分配 LUN。
- LU (logical unit) number (LU (逻辑单元) 号) – 将 LUN 与目标和启动器组相关联时，会按目标组和启动器对向 LUN 分配唯一的逻辑单元号。可由一个启动器通过一个目标组访问的两个 LUN 不能共享一个逻辑单元号。该属性控制逻辑单元是必须包含数字 0 还是自动分配的编号。
- Operational status (运行状态) – 此 LUN 的运行状态。无论目标或启动器的配置如何，启动器都无法访问脱机的 LUN。
- Write cache behavior (写入高速缓存行为) – 此设置控制 LUN 高速缓存是否写入。当此设置为关闭状态时，所有写入都是同步的，如果没有日志设备可用，则写入性能会大受影响。因此，启用此设置可以显著提高写入性能，但在意外关闭或故障转移时也会造成数据损坏，除非客户机应用程序了解可变写入高速缓存的语义并在必要时适当地刷新高速缓存。在启用此设置之前，请查阅客户机应用程序文档。
- GUID – LUN 的 GUID 是用于标识 SCSI 设备的全局唯一的只读标识符。此 GUID 在不同的机头节点和重复的环境中将保持一致。

共享资源 – HTTP

表 12-11 共享资源 – HTTP 属性

属性	说明
Share mode (共享模式)	此文件系统的 HTTP 共享模式。为 none (无)、read only (只读) 或 read/write (读/写) 之一。

共享资源 – FTP

表 12-12 共享资源 – FTP 属性

属性	说明
Share mode (共享模式)	此文件系统的 FTP 共享模式。为 none (无)、read only (只读) 或 read/write (读/写) 之一。

共享资源 – SFTP

表 12-13 共享资源 – SFTP 属性

属性	说明
Share mode (共享模式)	此文件系统的 SFTP 共享模式。为 none (无)、read only (只读) 或 read/write (读/写) 之一。

Shares (共享资源) > Shares (共享资源) > Access (访问)

访问控制

通过此视图，您可以设置选项来控制 ACL 行为并控制对文件系统根目录的访问。此视图仅适用于文件系统。

共享资源 – 根目录访问权限

控制文件系统根目录的基本访问控制。可以通过正在使用的任何协议对这些设置进行带内管理，但为方便起见，也可以在此处指定它们。在只读文件系统中，无法更改这些属性，因为它们需要更改文件系统根目录的元数据。

共享资源 – 用户

根目录的所有者。可将其指定为用户 ID 或用户名。有关映射 Unix 和 Windows 用户的更多信息，请参见“身份映射” [233]服务。对于基于 Unix 的 NFS 访问，这可以使用 chown 命令从客户机进行更改。

共享资源 – 组

根目录的组。可将其指定为组 ID 或组名称。有关映射 Unix 和 Windows 组的更多信息，请参见“[身份映射](#)” [233]服务。对于基于 Unix 的 NFS 访问，这可以使用 `chgrp` 命令从客户机进行更改。

共享资源 – 权限

对根目录的标准 Unix 权限。对于基于 Unix 的 NFS 访问，这可以使用 `chmod` 命令从客户机进行更改。这些权限分为三种类型。

表 12-14 共享资源用户

访问类型	说明
User (用户)	作为目录的当前所有者的用户。
Group (组)	作为目录的当前组的组。
Other (其他)	所有其他访问。

对于每个访问类型，可以授予以下权限。

表 12-15 共享资源权限

类型		说明
读取	R	列出目录内容的权限。
写入	W	在目录中创建文件的权限。*
执行	Trusted_Certificate.htm X	在目录中查找条目的权限。如果用户具有执行权限但没有读取权限，则他们可以通过名称显式访问文件，但无法列出目录的内容。

- 自 2011.1 软件发行版起，针对所有目录的“写入”权限关联了以下附加行为：
- 可以删除目录中的子文件（与 ACL D 权限相同），除非在目录上设置了 `sticky` 位，在这种情况下，只有当文件的所有者请求删除时才可以删除子文件
- 可以更改与文件或目录关联的时间（与 ACL A 权限相同）
- 可以创建扩展的属性，而且允许写入到扩展的属性目录（与 ACL W 权限相同）

在 BUI 中，通过单击各个框选择权限。或者，单击标签（“user”（用户）、“group”（组）或 “other”（其他））将选择（或取消选择）此标签中的所有权

限。在 CLI 中，权限以标准 Unix 八进制值指定，其中每个数字分别对应于（按顺序）"user"、"group" 和 "other"。每个数字均为读取 (4)、写入 (2) 和执行 (1) 三者之和。因此，权限值 743 等同于 "user" 的权限为 RWX、"group" 的权限为 R，"other" 的权限为 WX。

除了在共享资源创建时设置 POSIX 权限位，管理员还可以通过另一种方式设置权限，即选择 "Use Windows Default Permissions"（使用 Windows 默认权限）选项，此选项将应用 ACL（如下面的“[root directory ACL](#)”[297]一节所述）。这是一种简化管理的快捷方式，适用于完全或主要由具有 Windows 背景知识的用户管理的环境，这种方式旨在提供类似于在 Windows 服务器上创建共享资源的行为。

共享资源 – ACL 行为

有关 ACL 及其如何工作的信息，请参见“[root directory ACL](#)”[297]文档。

ACL Behavior on Mode Change（模式更改时的 ACL 行为）

如果使用标准 Unix user/group/other 权限通过 `chmod(2)` 修改 ACL，简化的模式更改请求与现有 ACL 进行交互时，其交互方式因该属性的设置而异。

表 12-16 模式更改值

BUI 值	CLI 值	说明
Discard ACL（放弃 ACL）	discard	不表示目录或文件模式的所有 ACL 条目都将被放弃。这是默认行为。
Mask ACL with mode（通过模式屏蔽 ACL）	mask	除非用户项与文件或目录的所有者具有相同的 UID，否则将减少权限，以使其不会大于组权限位。在这种情况下，减少 ACL 权限，以使其不会大于所有者权限位。如果未执行显式 ACL 集合操作，则 mask 值还会在模式更改之后保留 ACL。
Do not change ACL（不更改 ACL）	passthrough	除生成用于表示文件或目录的新模式的 ACL 条目外，不对 ACL 进行任何更改。

ACL Inheritance Behavior（ACL 继承行为）

创建新文件或目录时，可能会从父目录继承现有 ACL 设置。此属性控制此继承的方式。以下属性设置通常仅影响标记为可继承的 ACL 条目 – 不管对此属性做何设置，均不会

传播到其他条目。但是，与 SMB 结合使用时所有普通 ACL 条目都是可继承条目。普通 ACL 表示传统的 Unix owner/group/other 条目。

表 12-17 ACL 继承行为值

BUI 值	CLI 值	说明
Do not inherit entries (不继承条目)	discard	不继承任何 ACL 条目。根据所用客户机和协议创建文件或目录。
Only inherit deny entries (仅继承权限为 deny (拒绝) 的条目)	noallow	仅继承指定了 "deny" 权限的可继承 ACL 条目。
Inherit all but "write ACL" and "change owner" (全部继承但 "write ACL" (写入 ACL) 和 "change owner" (更改所有者) 除外)	restricted	继承 ACL 条目时删除 "write_acl" 和 "write_owner" 权限，但不更改可继承 ACL 条目的其他权限。这是默认设置
Inherit all entries (继承所有条目)	passthrough	继承所有可继承的 ACL 条目。"passthrough" 模式通常用于在目录树中使用相同模式创建所有"数据"文件。管理员设置 ACL 继承，使所有文件以某个特定模式 (例如 0664 或 0666) 创建。
Inherit all but "execute" when not specified (未指定时全部继承但 "execute" (执行) 除外)	passthrough-x	此行为与 "passthrough" 的行为基本相同，只是仅在文件创建模式还要求执行位时，owner、group 和 everyone ACL 条目才会继承执行权限。如果 "passthrough" 设置可以按预期处理数据文件，但是您希望将文件创建模式中的执行位选择性地包括到继承的 ACL 中，可以使用此设置。一个示例是通过工具生成的输出文件，例如 "cc" 或 "gcc"。如果继承的 ACL 不包括执行位，通过编译器生成的输出可执行文件将无法执行，除非使用 chmod(1) 更改此文件的权限。

使用 SMB 在具有普通 ACL 的目录中创建文件时，将继承所有 ACL 条目。因此，会出现以下行为：

- 继承位的显示会因在 SMB 还是 NFS 中查看而有所不同。在 SMB 中查看 ACL 目录时，会显示继承位。在 NFS 中，不会显示继承位。
- 使用 SMB 在目录中创建文件时，其 ACL 条目将显示为已继承；但是，通过 NFS 查看时，此目录显示不具有任何可继承的 ACL 条目。
- 如果 ACL 发生了更改，不再为普通条目 (例如通过添加访问控制条目 (access control entry, ACE))，不会出现此行为。
- 如果使用 SMB 修改了 ACL，则产生的 ACL 会将先前虚构的继承位转换为真实的继承位。

所有上述行为在未来发行版中都有可能更改。

根目录 ACL

通过访问控制列表管理对文件和目录的细粒度访问。ACL 描述是为特定用户或组授予的权限（如果有）。此 ZFSSA 支持 NFSv4 样式的 ACL，还可通过 SMB 访问 ACL。不支持 POSIX 草案 ACL（由 NFSv3 使用）。某些普通 ACL 可通过 NFSv3 表示，但是在通过 NFSv3 进行访问时，对 ACL 进行的复杂更改可能导致出现未定义的行为。

与根目录访问一样，此属性仅影响文件系统的根目录。ACL 可通过带内协议管理进行控制，但是 BUI 提供了一种仅针对文件系统的根目录设置 ACL 的方式。无法通过 CLI 设置根目录 ACL。如果没有 BUI 这种选择，则可以使用带内管理工具。更改此 ACL 不会影响文件系统现有的文件和目录。新创建的文件和目录可能继承、也可能不继承这些设置，具体取决于 ACL 继承行为。但是，使用 SMB 在具有普通 ACL 的目录中创建文件时，将继承所有 ACL 条目。

一个 ACL 由任意数量的 ACE（Access Control Entry，访问控制条目）组成。每个 ACE 描述了一个类型/目标、一个模式、一组权限和数个继承标志。ACE 按顺序（从 ACL 的开头开始）应用，以确定是否允许给定操作。有关通过数据协议对 ACL 进行带内配置的信息，请查阅相应的客户机文档。下面介绍了用于管理 ACL 的 BUI 界面以及对根目录的影响。

表 12-18 共享资源 – ACL 类型

类型	说明
所有者	目录的当前所有者。如果更改所有者，此 ACE 将应用于新所有者。
组	目录的当前组。如果更改组，此 ACE 将应用于新组。
所有人	任意用户。
已命名用户	由 "target"（目标）字段指定的用户。可以用用户 ID 或可由当前名称服务配置解析的名称来指定该用户。
已命名组	由 "target"（目标）字段指定的组。可以用组 ID 或可由当前名称服务配置解析的名称来指定该组。

表 12-19 共享资源 – ACL 模式

模式	说明
 Allow（允许）	显式授予 ACE 目标权限。
 Deny（拒绝）	显式拒绝授予 ACE 目标权限。

表 12-20 共享资源 – ACL 权限

	权限	说明
	读取	
(r)	Read Data/List Directory (读取数据/列出目录)	用于列出目录内容的权限。由文件继承时, 此权限允许读取文件的数据。
(x)	Execute File/Traverse Directory (执行文件/遍历目录)	此权限允许遍历 (查找) 目录中的条目。由文件继承时, 此权限允许执行文件。
(a)	Read Attributes (读取属性)	用于读取文件基本属性 (非 ACL) 的权限。基本属性视为静态级别属性, 授予此权限意味着用户可执行 ls 和 stat 等效项。
(R)	Read Extended Attributes (读取扩展属性)	此权限允许读取文件的扩展属性或查找扩展属性目录。
	写入	
(w)	Write Data/Add File (写入数据/添加文件)	用于将新文件添加到目录的权限。由文件继承时, 此权限允许在文件中的任意位置 (在文件偏移范围内) 修改数据。此权限包括增加文件大小或在任意偏移位置处写入的能力。
(p)	Append Data/Add Subdirectory (附加数据/添加子目录)	此权限允许在目录中创建子目录。由文件继承时, 此权限允许修改文件的数据 (仅允许从文件末尾开始修改)。当前不支持将此权限应用到文件。
(d)	Delete (删除)	用于删除文件的权限。
(D)	Delete Child (删除子项)	删除目录中的文件的权限。从 2011.1 软件发行版开始, 如果设置了粘滞位, 子文件只能由文件所有者删除。
(A)	Write Attributes (写入属性)	此权限允许更改与文件或目录关联的时间。
(W)	Write Extended Attributes (写入扩展属性)	用于创建扩展属性或写入扩展属性目录的权限。
	管理员	
(c)	Read ACL/Permissions (读取 ACL/权限)	读取 ACL 的权限。
(C)	Write ACL/Permissions (写入 ACL/权限)	此权限允许写入 ACL 或更改基本访问模式。
(o)	Change Owner (更改所有者)	此权限允许更改所有者。

	权限	说明
	Inheritance (继承)	
(f)	Apply to Files (应用于文件)	继承到目录中所有新建的文件。
(d)	Apply to Directories (应用于目录)	继承到目录中所有新建的目录。
(i)	Do not apply to self (不应用于自身)	当前 ACE 不应用到当前目录，但会应用到子项。此标志要求设置了 "Apply to Files" (应用于文件) 或 "Apply to Directories" (应用于目录)。
(n)	Do not apply past children (不应用过去的子项)	当前 ACE 应仅在树中的一个级别继承，即由直接子项继承。此标志要求设置了 "Apply to Files" (应用于文件) 或 "Apply to Directories" (应用于目录)。

在创建共享资源时如果选择了使用 Windows 默认权限的选项时，将为共享资源的根目录创建具有以下三个条目的 ACL：

表 12-21 共享资源根目录实体

类型	操作	访问权限
所有者	允许	完全控制
组	允许	读取和执行
所有人	允许	读取和执行

共享资源 – 快照

快照是文件系统在给定时间点的只读副本。有关快照及其工作原理的更多信息，请参见[“概念” \[262\]](#)页。

共享资源 – 快照属性

.zfs/snapshot visible (.zfs/snapshot 可见性)

可通过数据协议在文件系统根目录的 .zfs/snapshot 中访问文件系统快照。文件系统的所有快照的一个列表，可以像访问常规文件系统数据那样访问这些快照（以只读模式）。默认情况下，在列出目录内容时 ".zfs" 目录不可见，但可通过显式查找来访问该目录。隐藏该目录可以防止备份软件时不仅备份了新数据还无意中备份了快照。

表 12-22 快照值

BUI 值	CLI 值	说明
Hidden (隐藏)	hidden	在列出文件系统根目录的目录内容时，.zfs 目录不可见。这是默认设置。
Visible (可见)	visible	此 .zfs 目录像文件系统任何其他目录一样正常显示。

调度的快照标签

该可选属性会将用户定义的标签附加到每个调度的快照，默认情况下为空。标签可以针对单个共享资源设置，也可以针对项目设置而由其共享资源继承，但是不能同时针对两者设置。快照标签有助于标识创建了快照的项目或共享资源，例如 "project1:share1" 可以表示为 project1 中的 share1 创建了调度的快照。标签最多可以包含 35 个字母数字字符，并且可以包含特殊字符 _ - . :

使用 BUI 列出快照

"snapshots" (快照) 选项卡下是共享资源的活动快照列表。此列表分为两个选项卡："Snapshots" (快照) 选项卡用于浏览和管理快照。"Schedules" (调度表) 选项卡管理自动快照调度表。在 "Snapshots" (快照) 选项卡中，可以选择查看所有快照、仅手动快照或仅调度的快照。对于每个快照，将显示以下字段：

字段	说明
Name (名称)	快照的名称。包括两种类型的快照：手动和自动。 手动快照："Name" (名称) 是创建快照时提供的名称。通过单击该名称并输入新值可重命名手动快照。 自动快照：包括三种类型，不能对其进行重命名： - .auto：用户配置的调度快照，应用定制保留策略 (请参见 "Scheduled Snapshots" [303])。 - .ndmp：用于 NDMP 备份和自动删改。 - .rr：用于远程复制和自动删改。
Creation (创建)	创建快照的日期和时间。
Unique (唯一)	快照使用的唯一空间的大小。快照最初引用的所有块与文件系统或 LUN 自身相同。随着活动文件系统的分化，活动共享资源中已更改的块可能仍保存在一个或多个快照中。当一个块属于多个快照时，它将计入共享资源快照使用空间中，但不会显示在任何特定快

字段	说明
	照的唯一空间中。唯一空间包含仅保存在一个特定快照中的块，表示在销毁该快照时会释放的空间大小。
Total (总量)	快照引用的空间总量。该值表示获取快照时文件系统的大小，理论上，任何快照所占有的空间大小等于重写数据块时的总大小。
Clones (克隆)	显示快照的“克隆” [262] 数量。将鼠标悬停在克隆数量非零的快照行上，将出现“Show...”（显示...）链接。单击此链接将打开一个对话框，其中显示所有克隆的完整列表。

使用 BUI 创建手动快照

包括两种类型的快照：项目级和共享资源/LUN 级快照。

▼ 创建项目级快照

1. 打开要创建快照的项目。
2. 单击 "Snapshots"（快照）选项卡。
3. 单击  图标。此时将显示快照列表。
4. 在对话框中，键入快照的名称。
5. 要创建快照，请单击 "apply"（应用）。

▼ 创建共享资源/LUN 级快照

1. 打开要创建快照的共享资源/LUN。
2. 单击 "Snapshots"（快照）选项卡。
3. 单击  图标。此时将显示快照列表。
4. 在对话框中，键入快照的名称。
5. 要创建快照，请单击 "apply"（应用）。

对可以创建的快照数量没有限制，但是每个快照都会占用内存，因此创建大量的快照会降低系统运行速度。系统范围内快照数量的实际限制取决于系统配置，但应超过十万。

▼ 重命名快照 (BUI)

1. 要重命名快照，请单击活动快照列表中的名称。这将转到文本输入框。
2. 更新文本输入框中的名称后，按回车键或更改焦点均将提交更改。

▼ 销毁快照 (BUI)

1. 要销毁快照，请将鼠标悬停在目标快照的行上，单击  图标。
2. 销毁快照将需要销毁所有克隆及其后代。如果是这样的话，您将收到提示，其中显示会受影响的克隆的列表。

▼ 回滚到某个快照 (BUI)

1. 要回滚文件系统，请针对目标快照单击  图标。
2. 此时将显示确认对话框，如果存在任何快照克隆、任何较新的快照或其后代，它们均会显示出来，并指出它们将在此过程中销毁。

除访问文件系统快照目录中的数据外，快照还可用于回滚到文件系统或 LUN 以前的实例。这需要销毁所有较新的快照及其克隆，并将共享资源内容恢复到创建此快照时的样子。尽管对文件系统根目录访问权限的更改将会丢失（因为它是文件系统数据的一部分），但不会影响共享资源的任何属性设置。

▼ 克隆快照 (BUI)

- 要创建克隆，请针对源快照单击  图标。对话框将提示以下值。
 - Project（项目）– 目标项目。默认情况下，在当前项目中创建克隆，但是也可以在其他项目中创建克隆（或在项目之间移动克隆）。
 - Name（名称）– 为克隆键入名称。

-
- Mountpoint (多点) – 要使用此值, 请单击锁定图标。为克隆设置多点。设置 "Retain Other Local Settings" (保留其他本地设置) 时, 必须为克隆指定不同的挂载点, 因为共享资源无法保存相同的挂载点。
 - Resource Name (资源名称) – 要使用此值, 请单击锁定图标。输入您希望用于克隆的资源。
 - Retain Other Local Settings (保留其他本地设置) – 默认情况下, 克隆中当前所有继承的文件系统属性将从目标项目继承。始终保留本地设置。设置此属性会使所有继承属性在新克隆中保留为本地设置。

“克隆” [262] 是快照的可写副本, 可以像任何其他共享资源那样进行管理。与文件系统的快照一样, 它最初不占用任何额外空间。随着克隆中的数据发生变化, 克隆占用的空间将逐渐增加。如果不销毁克隆, 原始快照无法销毁。调度的快照可以安全克隆, 销毁时会忽略具有克隆的调度快照。

使用 BUI 调度的快照

除手动快照外, 还可以根据下表配置自动快照。这些快照以 ".auto-<timestamp>" 格式命名, 可以按每半小时、每小时、每天、每周或每月为时间间隔创建。调度表是时间间隔和保留策略的列表。

显示的时间采用本地 (客户机浏览器) 时区。但是, 时间以 UTC 格式存储和执行, 而不考虑夏令时之类的约定。例如, 安排在 PST (UTC-8) 时间上午 10:00 的快照将以 UTC 时间存储并在 UTC 时间 18:00 存储和执行。

可以对一个项目或一个共享资源设置自动快照, 但不能同时对这两者设置自动快照。否则, 重叠的调度表和保留策略会导致无法保证这两个调度表。删除一个时间间隔或更改保留策略均将立即销毁新调度表未涵盖的所有自动快照。将忽略具有克隆的自动快照。

以前版本的软件允许以分钟为频率创建自动快照。这已证明会给系统带来额外负担且通常没有用处。为了帮助用户避免给系统带来额外负担, 在 2010.Q3 发行版中已删除此功能。现在只能以半小时或更长时间为单位指定快照期间。将保留现有分钟期间以备软件回滚, 以前的实例将根据现有调度表过期, 但不会创建同类的新快照。如果系统发现使用此频率的共享资源或项目, 将发出警报。



- 要添加新时间间隔, 请在查看 "Schedules" (调度表) 选项卡时单击  图标。每个时间间隔具有以下属性。

属性	说明
Frequency (频率)	"half hour" (半小时)、"hour" (一小时)、"day" (一天)、"week" (一周) 或 "month" (一月) 其中之一这表示创建快照的频率。
Offset (偏移)	该属性指定频率内的偏移量。例如，选择小时频率时，可以明确指定小时内的分钟偏移量来创建快照。对于每天快照，可以按小时和分钟指定偏移量，对于每周或每月快照，可以按天、小时和分钟指定偏移量。
Keep at most (保存最多)	控制快照的保留策略。自动快照可以永久保留（半小时和小时快照除外，这些快照的上限分别是 48 和 24 小时）或者限定为一定数量。如果给定时间间隔的自动快照早于保留策略，此限制将导致这些快照被删除。这实际由获取这些快照的时间控制，而不是由绝对数控制。因此，如果您采用小时快照而 ZFSSA 停机一天，则在 ZFSSA 重新运行后，所有小时快照都将被删除。对于属于多个时间间隔的快照，仅当没有任何时间间隔指定应保留该快照时，才会将其销毁。

使用 CLI 创建手动快照

要访问共享资源的快照，请导航到该共享资源和快照上下文。

```
clownfish:> shares select default select builds
clownfish:shares default/builds> snapshots
clownfish:shares default/builds snapshots>
```

列出快照 (CLI)

可使用标准 CLI 命令列出快照。

```
clownfish:shares default/builds snapshots> list
today
yesterday
clownfish:shares default/builds snapshots>
```

创建手动快照 (CLI)

要创建手动项目级快照，请导航到项目和快照节点，然后使用 snapshot 命令：

```
clownfish:cd /
clownfish:shares select myproject snapshots
clownfish:shares myproject snapshots> snapshot cob_monday
```

要为单个共享资源创建手动共享资源级快照，请导航到该共享资源，在此处使用 `snapshot` 命令：

```
clownfish:cd /
clownfish:shares select myproject select share1 snapshots
clownfish:snapshot lunchtime
```

重命名快照 (CLI)

要重命名快照，请使用 `rename` 命令：

```
clownfish:shares default/builds snapshots> rename test test2
clownfish:shares default/builds snapshots>
```

销毁快照 (CLI)

要销毁快照，请使用 `destroy` 命令：

```
clownfish:shares default/builds snapshots> select test2
clownfish:shares default/builds@test2> destroy
This will destroy this snapshot. Are you sure? (Y/N)
clownfish:shares default/builds snapshots>
```

您还可以在共享资源上下文中使用 `destroy` 命令而无需选择单个快照：

```
clownfish:shares default/builds snapshots> destroy test2
This will destroy this snapshot. Are you sure? (Y/N)
clownfish:shares default/builds snapshots>
```

回滚到某个快照 (CLI)

要回滚到某个快照，请选择目标快照并运行 `rollback` 命令：

```
clownfish:shares default/builds snapshots> select today
clownfish:shares default/builds@today> rollback
Rolling back will revert data to snapshot, destroying newer data. Active
initiators will be disconnected.

Continue? (Y/N)
```

```
clownfish:shares default/builds@today>
```

克隆快照 (CLI)

要克隆快照，请使用 `clone` 命令。此命令将您置于一个未提交的共享资源上下文中，该上下文与用于创建共享资源的上下文相同。在此处，您可以根据需要调整属性，然后再提交更改以创建克隆。

```
clownfish:shares default/builds snapshots> select today
clownfish:shares default/builds@today> clone testbed
clownfish:shares default/testbed (uncommitted clone)> get
    aclinherit = restricted (inherited)
    aclmode = discard (inherited)
    atime = true (inherited)
    checksum = Fletcher4 (inherited)
    compression = off (inherited)
    copies = 1 (inherited)
    mountpoint = /export/testbed (inherited)
    quota = 0 (default)
    readonly = false (inherited)
    recordsize = 128K (inherited)
    reservation = 0 (default)
    secondarycache = all (inherited)
    nbmand = false (inherited)
    sharesmb = off (inherited)
    sharenfs = on (inherited)
    snapdir = hidden (inherited)
    vscan = false (inherited)
    sharedav = off (inherited)
    shareftp = off (inherited)
    root_group = other (default)
    root_permissions = 777 (default)
    root_user = nobody (default)
    quota_snap = true (default)
    reservation_snap = true (default)
clownfish:shares default/testbed (uncommitted clone)> set quota=10G
    quota = 10G (uncommitted)
clownfish:shares default/testbed (uncommitted clone)> commit
clownfish:shares default/builds@today>
```

此命令还支持一个可选参数（第一个参数），此参数是将在其中创建克隆的项目。默认情况下，将在要克隆的共享资源所属的项目中创建克隆。

使用 CLI 列出从属克隆

要列出从特殊快照（从属克隆）中创建的所有克隆，请导航至该快照，然后使用列出克隆命令。

```
clownfish:shares default/builds> snapshots
```

```
clonefish:shares default/builds snapshots> select today
clonefish:shares default/builds@today> list clones
```

```
Clones: 2 total
```

```
PROJECT      SHARE
default      testbed
default      production
clonefish:shares default/builds@today>
```

结果显示克隆名称和克隆所在的项目。

使用 CLI 调度的快照

可以在单个共享资源的项目级使用 `automatic` 命令在快照上下文中配置自动调度的快照。在此上下文中，可以使用 `create` 和 `destroy` 命令添加和删除新时间间隔。每个时间间隔都具有一组属性，这些属性与 BUI 视图中的频率、偏移量和要保留的快照数相对应。调度表是以 UTC 格式维护的。

```
clonefish:shares default/builds snapshots> automatic
clonefish:shares default/builds snapshots automatic> create
clonefish:shares default/builds snapshots automatic (uncommitted)> set frequency=day
    frequency = day (uncommitted)
clonefish:shares default/builds snapshots automatic (uncommitted)> set hour=14
    hour = 14 (uncommitted)
clonefish:shares default/builds snapshots automatic (uncommitted)> set minute=30
    minute = 30 (uncommitted)
clonefish:shares default/builds snapshots automatic (uncommitted)> set keep=7
    keep = 7 (uncommitted)
clonefish:shares default/builds snapshots automatic (uncommitted)> get
    frequency = day (uncommitted)
    day = (unset)
    hour = 14 (uncommitted)
    minute = 30 (uncommitted)
    keep = 7 (uncommitted)
clonefish:shares default/builds snapshots automatic (uncommitted)> commit
clonefish:shares default/builds snapshots automatic> list
NAME          FREQUENCY    DAY          HH:MM KEEP
automatic-000 day          -            14:30    7
clonefish:shares default/builds snapshots automatic> done
clonefish:shares default/builds snapshots>
```

使用 CLI 设置调度的快照标签

在 BUI 中，可以对项目或共享资源设置 "scheduled snapshot label"（调度的快照标签）属性。同样，在 CLI 中，可以通过首先导航到项目上下文或共享资源上下文来设置标签。要创建调度的快照标签，请使用 `set snaplabel` 命令：

```
clonefish:shares project1/share1> set snaplabel=project1:share1
```

项目

共享资源、文件系统和 LUN 都可以组合到项目中。项目定义用于管理共享资源的常见管理控制点。项目内的共享资源可以共享通用设置，并且除了在共享资源级别之外还可以在项目级别强制实施配额。项目还可以仅用于将逻辑相关的共享资源组合在一起，从而可以从单个点访问其通用属性（例如累积空间）。

默认情况下，第一次配置存储池时 ZFSSA 创建一个默认项目。可以在该默认项目内创建所有共享资源，但是对于规模适中的环境，强烈建议创建其他项目（如果仅用于组织目的）。

通过 BUI 使用项目

通过 "Shares"（共享资源）-> "Projects"（项目），可访问 "Projects"（项目）UI。此 UI 显示了系统中所有项目的列表，不过，也可以通过使用项目面板或在项目中编辑共享资源时单击项目名称来选择项目。

项目字段

导航到项目视图后，该视图将显示系统中项目的列表。或者，您可以导航到共享资源屏幕并打开项目面板，以快捷地访问项目。此面板不能充分扩展以容纳大量的项目且不能代替完整项目列表。每个项目均包含以下字段：

表 12-23 项目字段

字段	说明
Name（名称）	共享资源的名称。共享资源名称是可编辑的文本字段。单击此名称可以为项目输入新名称。按回车键或从名称移动焦点将会提交更改。系统将要求您确认该项操作，因为对共享资源进行重命名需要断开活动客户端机的连接。
Size（大小）	项目中的所有共享资源以及未使用的预留空间的总大小。

以下工具可用于每个项目：

表 12-24 项目图标

图标	说明
	编辑单个项目（也可通过双击行来编辑）。

图标	说明
	销毁项目。系统将会提示您确认此项操作，因为这将销毁共享资源中的所有数据，并且无法撤消。

编辑项目

要编辑某项目，请单击铅笔图标或双击项目列表中的行，或者单击项目面板中的名称。此操作将选择该项目，并显示多个不同的选项卡供选择来编辑项目的属性。

项目的名称显示在项目面板右侧的左上角。也可以通过单击项目名称并在输入框中输入新文本来更改项目的名称。系统将要求您确认该项操作，因为这将需要断开项目的活动客户机的连接。

使用情况统计信息

视图的左侧（展开时位于项目面板下面）是一个表，用于说明当前空间使用情况统计信息。如果任何属性为零，则表中将排除该属性。其中的大部分属性是项目和共享资源所共有的，但某些统计信息仅对项目有意义。

- Available space（可用空间） – 请参见[“Shares（共享资源）> Shares（共享资源）”](#) [273]。
- Referenced data（引用数据） – 项目中所有共享资源的所有引用数据与项目开销所占的少量空间之和。有关如何计算共享资源的引用数据的详细信息，请参见[“Shares（共享资源）> Shares（共享资源）”](#) [273]。
- Snapshot data（快照数据） – 所有共享资源的所有快照数据以及任何项目快照开销之和。有关如何计算共享资源的快照数据的详细信息，请参见[“Shares（共享资源）> Shares（共享资源）”](#) [273]。
- Unused Reservation（未使用的预留空间） – 项目的未使用预留空间。这仅包括项目级别预留空间中当前未使用的部分。不包括项目所含任何共享资源的未使用预留空间。
- Unused Reservation of shares（共享资源的未使用预留空间） – 所有共享资源的未使用预留空间之和。有关如何计算共享资源的未使用预留空间的详细信息，请参见[“Shares（共享资源）> Shares（共享资源）”](#) [273]。
- Total space（总空间） – 引用数据、快照数据、未使用预留空间和共享资源的未使用预留空间之和。

静态属性

编辑特定项目时，共享资源视图的左侧还显示静态属性。这些属性为只读属性，不能修改。

- Compression ratio（压缩率） – 有关完整说明，请参见[“Shares（共享资源）> Shares（共享资源）”](#) [273]。

▼ 创建项目

1. 要创建项目，请查看项目列表并单击  按钮。
2. 或者，单击项目面板中的 "Add..." (添加...) 按钮将显示相同的对话框。输入项目名称并单击 "Apply" (应用) 以创建项目。

通过 CLI 使用项目

项目 CLI 位于 shares 下

导航

要选择项目，请使用 select 命令：

```
clownfish:> shares
clownfish:shares> select default
clownfish:shares default> get
      aclinherit = restricted
      aclmode = discard
      atime = true
      checksum = fletcher4
      compression = off
      compressratio = 100
      copies = 1
      creation = Thu Oct 23 2009 17:30:55 GMT+0000 (UTC)
      mountpoint = /export
      quota = 0
      readonly = false
      recordsize = 128K
      reservation = 0
      secondarycache = all
      nbmand = false
      sharesmb = off
      sharenfs = on
      snapdir = hidden
      snaplabel = project1:share1
      vscan = false
      sharedav = off
      shareftp = off
      default_group = other
      default_permissions = 700
      default_sparse = false
      default_user = nobody
      default_volblocksize = 8K
      default_volsize = 0
      space_data = 43.9K
      space_unused_res = 0
```

```

space_unused_res_shares = 0
space_snapshots = 0
space_available = 12.0T
space_total = 43.9K
clownfish:shares default>

```

项目操作

可以使用 `project` 命令创建项目。在提交更改之前，可以根据需要对属性进行修改：

```

clownfish:shares> project home
clownfish:shares home (uncommitted)> get
    mountpoint = /export (default)
    quota = 0 (default)
    reservation = 0 (default)
    sharesmb = off (default)
    sharenfs = on (default)
    sharedav = off (default)
    shareftp = off (default)
    default_group = other (default)
    default_permissions = 700 (default)
    default_sparse = true (default)
    default_user = nobody (default)
    default_volblocksize = 8K (default)
    default_volsize = 0 (default)
    aclinherit = (default)
    aclmode = (default)
    atime = (default)
    checksum = (default)
    compression = (default)
    copies = (default)
    readonly = (default)
    recordsize = (default)
    secondarycache = (default)
    nbmand = (default)
    snapdir = (default)
    snaplabel = project1:share1
    vscan = (default)
    custom:contact = (default)
    custom:department = (default)
clownfish:shares home (uncommitted)> set sharenfs=off
    sharenfs = off (uncommitted)
clownfish:shares home (uncommitted)> commit
clownfish:shares>

```

可以使用 `destroy` 命令销毁项目：

```

clownfish:shares> destroy home
This will destroy all data in "home"! Are you sure? (Y/N)
clownfish:shares>

```

选择项目之后，还可以在项目上下文中运行此命令。

可以使用 `rename` 命令重命名项目：

```
clownfish:shares> rename default home
clownfish:shares>
```

选择群集中的池

在主动/主动群集配置中，故障转移时，一个节点可由两个池控制。在这种情况下，CLI 上下文将在括号中显示当前池。您可以在顶级共享资源上下文中使用 `set` 命令更改池：

```
clownfish:shares (pool-0)> set pool=pool-1
clownfish:shares (pool-1)>
```

选择池上下文后，将使用标准 CLI 界面管理该池中的项目和共享资源。

项目属性

下面显示了在 CLI 中提供的属性及其在 BUI 中对应的属性。可以使用标准 CLI 命令 `get` 和 `set` 来设置这些属性。此外，还可以使用 `unset` 命令从父项目中继承这些属性。

CLI 名称	“类型” [262]	BUI 名称	BUI 位置
aclinherit	继承	“项目访问” [320]	Access（访问）
aclmode	继承	“项目访问” [320]	Access（访问）
atime	继承	“项目 - 常规” [317]	General（常规）
checksum	继承	“项目 - 常规” [317]	General（常规）
compression	继承	“项目 - 常规” [317]	General（常规）
compressratio	只读	“项目” [312]	Static（静态）
copies	继承	“项目 - 常规” [317]	General（常规）
creation	只读	-	-
dedup	继承	“项目 - 常规” [317]	General（常规）
default_group	创建默认值	“项目 - 常规” [317]	General（常规）
default_permissions	创建默认值	“项目 - 常规” [317]	General（常规）
default_sparse	创建默认值	“项目 - 常规” [317]	General（常规）
default_user	创建默认值	“项目 - 常规” [317]	General（常规）
default_volblocksize	创建默认值	“项目 - 常规” [317]	General（常规）

CLI 名称	“类型” [262]	BUI 名称	BUI 位置
default_volsize	创建默认值	“项目 – 常规” [317]	General (常规)
mountpoint	继承	“项目 – 常规” [317]	General (常规)
nbmand	继承	“项目 – 常规” [317]	General (常规)
quota	空间管理	“项目 – 常规” [317]	General (常规)
readonly	继承	“项目 – 常规” [317]	General (常规)
recordsize	继承	“项目 – 常规” [317]	General (常规)
reservation	空间管理	“项目 – 常规” [317]	General (常规)
secondary cache	继承	“项目 – 常规” [317]	General (常规)
sharedav	继承	“项目协议” [319]	Protocols (协议)
shareftp	继承	“项目协议” [319]	Protocols (协议)
sharenfs	继承	“项目协议” [319]	Protocols (协议)
sharesmb	继承	“项目协议” [319]	Protocols (协议)
snapdir	继承	“项目快照” [320]	Snapshots (快照)
snaplabel	继承	“项目快照” [320]	Snapshots (快照)
space_available	只读	“项目” [312]	使用情况
space_data	只读	“项目” [312]	使用情况
space_snapshots	只读	“项目” [312]	使用情况
space_total	只读	“项目” [312]	使用情况
space_unused_res	只读	“项目” [312]	使用情况
space_unused_res_shares	只读	“项目” [312]	使用情况
vscan	继承	“项目 – 常规” [317]	General (常规)

项目 – 常规

项目 – 常规属性

在 BUI 中，这部分内容用于控制项目的整体设置，这些设置与任何特定协议无关，也与访问控制或快照无关。虽然 CLI 将所有属性组合到了单个列表中，但本节介绍了在这两种上下文中属性的行为。

有关如何将这些属性映射到 CLI 的信息，请参见[“项目 CLI”](#)。

项目 – 空间使用情况

存储池中的空间供所有共享资源使用。尽管可以以共享资源为单位实施空间限制，但文件系统可以根据需要动态地扩展或收缩。有关池存储的更多信息，请参见[“概念” \[262\]](#)页。

项目 – 配额

设置项目中所有文件系统和 LUN 所占用的总空间量的上限。有关更多信息，请参见[“共享资源” \[284\]](#)一节。与文件系统不同，项目配额不能排除快照，且只能在所有共享资源及其快照中执行。

项目 – 预留空间

为确保有空间可供项目中所有文件系统和 LUN 使用而预留的最小空间量。有关更多信息，请参见[“共享资源” \[284\]](#)一节。与文件系统不同，项目预留空间不能排除快照，且只能在所有共享资源及其快照中执行。

项目 – 继承的属性

这些属性是可由项目中的共享资源继承的标准属性。这些属性的行为与共享资源级别属性的行为相同，详细说明位于共享资源一节。

- [“挂载点” \[284\]](#)
- [“只读” \[284\]](#)
- [“Update access time on read \(读取时更新访问时间\)” \[284\]](#)
- [“Non-blocking mandatory locking \(非阻塞强制性锁定\)” \[284\]](#)
- [“Data compression \(数据压缩\)” \[284\]](#)
- [“Data deduplication \(重复数据删除\)” \[284\]](#)
- [“校验和” \[284\]](#)
- [“Cache device usage \(高速缓存设备使用\)” \[284\]](#)
- [“Database record size \(数据库记录大小\)” \[284\]](#)
- [“Additional replication \(额外复制\)” \[284\]](#)
- [“Virus scan \(病毒扫描\)” \[284\]](#)

项目 – 定制属性

可以根据需要添加定制属性，以便将用户定义的标记附加到项目和共享资源。有关更多信息，请参见[“模式” \[322\]](#)。

文件系统创建默认值

这些设置用于在创建文件系统时填充默认值。对其进行更改不会影响现有文件系统。在共享资源一节的相应部分可找到更多信息。

- “用户” [297]
- “组” [297]
- “权限” [297]

LUN 创建默认值

这些设置用于在创建 LUN 时填充默认值。对其进行更改不会影响现有 LUN。在共享资源一节的相应部分可找到更多信息。

- “卷大小” [284]
- “Thin provisioned (瘦置备)” [284]
- “Shares (共享资源) > Shares (共享资源)” [273]

项目协议

每个项目均具有特定于协议的属性，这些属性定义该项目中的共享资源在不同协议下的行为。通常，“共享资源” [290] 直接继承特定于协议的属性。例外和特殊情况如下。

- NFS – “NFS” [185] 共享资源属性正常继承，这在“共享资源文档” [290] 中进行了说明。
- SMB
 - Resource name (资源名称) – “SMB” [191] 客户机用来引用此共享资源的名称。
 - Enable Access-based Enumeration (启用基于访问的枚举) – 当启用了此选项时，将执行基于访问权限的枚举。基于访问权限的枚举根据客户机的凭证过滤目录条目。如果客户机无权访问某个文件或目录，则返回到客户机的条目列表中省略该文件或目录。默认情况下不启用此选项。

同一系统上的两个“SMB” [191] 共享资源不可具有相同的资源名称。文件系统从项目继承资源名称时，将根据以下规则构建共享资源的资源名称：

 - off – 所包含的文件系统不通过“SMB” [191] 导出。
 - on – 所包含的文件系统通过“SMB” [191] 导出，且其文件系统名称作为资源名称。
 - “off” 或 “on” 以外的其他任何值 – 采用 *<project's resource name>_<filesystem name>* 格式为每个文件系统构建资源名。
- iSCSI – 不继承“iSCSI” [189] 属性。

- HTTP – “HTTP” [208]共享资源属性正常继承，这在“共享资源文档” [290]中进行了说明。
- FTP – “FTP” [206]共享资源属性正常继承，这在“共享资源文档” [290]中进行了说明。
- SFTP – “SFTP” [216]共享资源属性正常继承，这在“共享资源文档” [290]中进行了说明。
- NFS – “NFS” [185]共享资源属性正常继承，这在“共享资源文档” [290]中进行了说明。
- SMB
 - Resource name（资源名称） – “SMB” [191]客户机用来引用此共享资源的名称。
 - Enable Access-based Enumeration（启用基于访问的枚举） – 当启用了此选项时，将执行基于访问权限的枚举。基于访问权限的枚举根据客户机的凭证过滤目录条目。如果客户机无权访问某个文件或目录，则返回到客户机的条目列表中并将省略该文件或目录。默认情况下不启用此选项。
同一系统上的两个“SMB” [191]共享资源不可具有相同的资源名称。文件系统从项目继承资源名称时，将根据以下规则构建共享资源的资源名称：
 - Off – 所包含的文件系统不通过“SMB” [191]导出。
 - On – 所包含的文件系统通过“SMB” [191]导出，且其文件系统名称作为资源名称。
 - "off" 或 "on" 以外的其他任何值 – 采用 *<project's resource name>_<filesystem name>* 格式为每个文件系统构建资源名。
- iSCSI – 不继承“iSCSI” [189]属性。

项目访问

- Access Control（访问控制） – 通过此视图可控制影响“ACL” [297]行为的可继承属性。
- Inherited ACL Behavior（继承的 ACL 行为） – 这些属性的行为方式与在共享资源级别相同。更改这些属性将更改当前继承这些属性的所有文件系统的相应行为。
 - “ACL behavior on mode change（模式更改时的 ACL 行为）” [297]
 - “ACL inheritance behavior（ACL 继承行为）” [297]

项目快照

快照是文件系统在给定时间点的只读副本。有关快照及其工作原理的更多信息，请参见“概念” [262]页。项目快照由项目中的每个文件系统和 LUN 的快照组成，这些快照都具有相同的名称。对于共享资源，可单独删除快照，并可使用与项目快照相同的名称创建共享资源快照（如果支持），但是这样做会导致未定义的行为，因为该快照会被视为具有相同名称的项目快照的一部分。

项目快照属性

.zfs/snapshot visible (.zfs/snapshot 可见性)

可通过数据协议在文件系统根目录的 .zfs/snapshot 中访问文件系统快照。文件系统的所有快照的一个列表，可以像访问常规文件系统数据那样访问这些快照（以只读模式）。默认情况下，在列出目录内容时 ".zfs" 目录不可见，但可通过显式查找来访问该目录。隐藏该目录可以防止备份软件时不仅备份了新数据还无意中备份了快照。

表 12-25 项目快照值

BUI 值	CLI 值	说明
Hidden (隐藏)	hidden	在列出文件系统根目录的目录内容时，.zfs 目录不可见。这是默认设置。
Visible (可见)	visible	此 .zfs 目录像文件系统任何其他目录一样正常显示。

调度的快照标签

该可选属性会将用户定义的标签附加到每个调度的快照，默认情况下为空。标签可以针对单个共享资源设置，也可以针对项目设置而由其共享资源继承，但是不能同时针对两者设置。快照标签有助于标识创建了快照的项目或共享资源，例如 "project1:share1" 可以表示为 project1 中的 share1 创建了调度的快照。标签最多可以包含 35 个字母数字字符，并且可以包含特殊字符 _ - . :

管理项目级别快照的方式与管理共享资源级别快照的方式相同。有关快照的更多信息，请参见[“共享资源：快照” \[303\]](#)。

项目快照不支持回滚或克隆操作。有关快照的更多信息，请参见[“共享资源：快照” \[303\]](#)。

要访问项目的快照，请导航到该项目并运行 snapshots 命令。

```
clownfish:> shares select default
clownfish:shares default> snapshots
clownfish:shares default snapshots>
```

在此之后，管理快照的方式与管理共享资源级别快照的方式相同。有关快照的更多信息，请参见[“共享资源：快照” \[303\]](#)。

项目快照不支持回滚或克隆操作。有关快照的更多信息，请参见[“共享资源：快照” \[303\]](#)。

模式

定制的共享资源属性

除了标准的内置属性之外，还可以配置可在所有共享资源和项目上使用的任意数目的其他属性。可以为这些属性指定基本类型以用于验证目的，并且它们可以与大多数其他标准属性一样被继承。这些值从不会被软件以任何方式使用，仅供最终用户使用。属性模式对系统来说是全局性的，在所有池之间保持一致，并且在群集对等设备之间保持同步。

在 BUI 中使用模式

要定义定制属性，请访问 "shares"（共享资源） > "schema"（模式）导航项。当前模式显示为一个列表，且可根据需要添加或删除条目。每个属性均包含以下字段：

表 12-26 模式属性字段

字段	说明
NAME（名称）	此属性的 CLI 名称。这只能包含字母数字字符或字符 "._:"。
DESCRIPTION（描述）	此属性的 BUI 名称。这可以包含任意字符，并且用于 CLI 的帮助部分中。
TYPE（类型）	属性类型，用于验证目的。这必须是下文介绍的类型之一。

属性的有效类型如下

表 12-27 属性的有效类型

BUI 类型	CLI 类型	说明
字符串	字符串	任意字符串数据。这等效于不进行验证。
Integer（整数）	Integer（整数）	一个正整数或负整数
Positive Integer（正整数）	PositiveInteger	正整数
Boolean（布尔）	Boolean	true/false 值。在 BUI 中这显示为一个复选框，而在 CLI 中该值必须是 "true" 或 "false" 值之一。
Email Address（电子邮件地址）	EmailAddress	一个电子邮件地址。仅执行最少的语法验证。

BUI 类型	CLI 类型	说明
Hostname 或 IP	Host	有效的 DNS 主机名或 IP (v4 或 v6) 地址。

按照属性表中提供的说明进行定义后，属性将出现在“[general](#)”（常规）”[284]属性选项卡下。属性由其 CLI 名称予以标识，因此重命名属性将具有删除系统上的所有现有设置的效果。被删除且之后又被重命名回原始名称的属性将仍然引用先前设置的值。更改属性的类型（支持该操作时）可能会对系统上的现有属性产生不确定的结果。现有属性将保留其当前设置，即使这些属性在新的属性类型下将无效也是如此。

▼ 使用 BUI 配置模式

- 1. 导航到 "Shares"(共享资源) > "Schema"（模式）视图
- 2. 单击 "+" 图标向模式属性列表添加一个新属性
- 3. 输入属性的名称 ("contact")
- 4. 输入属性的说明 ("Owner Contact")
- 5. 为新属性选择一个类型 ("Email Address")
- 6. 单击 "Apply"（应用）按钮
- 7. 导航到现有的共享资源或项目
- 8. 在 "Custom Properties"（定制属性）区域下更改 "Owner Contact"（所有者联系信息）属性。

通过 CLI 使用模式

可以在 "shares"（共享资源）> "schema"（模式）下找到模式上下文

```
carp:> shares schema
carp:shares schema> show
Properties:

NAME          TYPE          DESCRIPTION
owner         EmailAddress  Owner Contact
```

每个属性都是模式上下文的子代，使用属性名称作为标记。要创建属性，请使用 create 命令：

```
carp:shares schema> create department
carp:shares schema department (uncommitted)> get
      type = String
      description = department
carp:shares schema department (uncommitted)> set description="Department Code"
      description = Department Code (uncommitted)
carp:shares schema department (uncommitted)> commit
carp:shares schema>
```

在特定属性的上下文中，可使用标准 CLI 命令设置字段：

```
carp:shares schema> select owner
carp:shares schema owner> get
      type = EmailAddress
      description = Owner Contact
carp:shares schema owner> set description="Owner Contact Email"
      description = Owner Contact Email (uncommitted)
carp:shares schema owner> commit
```

定义了定制属性后，可以像访问任何名称为 "custom:<property>":

```
carp:shares default> get
...
      custom:department = 123-45-6789
      custom:owner =
...
carp:shares default> set custom:owner=bob@corp
      custom:owner = bob@corp (uncommitted)
carp:shares default> commit
```

▼ 使用 CLI 配置模式

1. 导航到模式上下文 (**shares schema**)
2. 创建一个名为 "contact" 的新属性 (**create contact**)
3. 设置属性的说明 (**set description="Owner Contact"**)
4. 设置属性的类型 (**set type=EmailAddress**)
5. 提交更改 (**commit**)
6. 导航到现有的共享资源或项目
7. 设置 "custom:contact" 属性

复制

许可声明：可以免费试用远程复制和克隆特性，但若要在生产环境中使用，必须为各特性购买单独的许可证。在试用期后，必须为这些特性获取许可证或停用它们。Oracle 保留随时审计许可合规性的权利。有关详细信息，请参阅 “Oracle Software License Agreement (“SLA”) and Entitlement for Hardware Systems with Integrated Software Options”。

复制概述

Oracle ZFS Storage Appliance 支持从源 ZFSSA 到任意数量的目标 ZFSSA 进行基于快照的项目和共享资源复制，可通过手动复制、按调度表复制或连续复制等方式进行。复制包括数据和元数据。远程复制（或“复制”）是满足一般用途的功能，该功能针对以下使用案例进行了优化：

- 灾难恢复。复制可用于镜像 ZFSSA，以供灾难恢复。如果发生影响主要 ZFSSA（甚至是整个数据中心）的服务的灾难，管理员可在灾难恢复站点激活服务，该站点将使用最近复制的数据接管工作。当主站点恢复后，在灾难恢复站点提供服务时更改的数据可以迁移回主站点并且将恢复正常服务。在发生此类灾难之前，完全可测试此类情况。
- 数据分发。在目标 ZFSSA 的客户机无法正常直接到达源 ZFSSA 的情况下，或这样的设置将产生过高的延迟时，可以使用复制将数据（例如虚拟机映像或媒体）分发到世界各地的远程系统上。例如，对本地缓存使用此方案以缩短只读数据（例如文档）的延迟。
- 磁盘到磁盘备份。对于磁带备份不可行的环境，可以使用复制作为备份解决方案。磁带备份可能不可行，例如，由于可用的带宽不足或进行恢复时延迟过高。
- 数据迁移。升级硬件或重新平衡存储时，可以使用复制在 ZFSSA 之间迁移数据和配置。也可以使用影子迁移来实现此目的。

远程复制功能具有以下几个重要属性：

- 基于快照。复制子系统在每次更新操作过程中都会创建一个快照。对于完整更新，将发送直到创建快照时的完整项目内容。对于增量更新，只发送自同一操作的上一个复制快照以后发生的更改。
- 块级别。每个更新操作将在块级别遍历文件系统，并将相应的文件系统数据和元数据发送到目标中。

- 异步。因为复制将创建快照并发送快照，所以在复制开始发送快照之前，有必要将数据提交到一个稳定的存储中。连续复制可有效地发送连续的文件系统更改流，但是对于 NAS 和 SAN 客户机，它仍然是异步的。
- 包括元数据。底层复制流将对用户数据和 ZFS 元数据进行序列化，包括在 "Shares"（共享资源）屏幕上配置的大多数属性。在首次复制更新完成后，可以在目标上修改这些属性，虽然在断开复制连接之前，不是所有修改都会生效。例如，这允许通过 NFS 共享到除源上的主机集以外的其他主机集上。有关详细信息，请参见[“管理复制数据包” \[335\]](#)。
- 安全。在 ZFS Storage Appliance 间使用的复制控制协议由 SSL 提供安全保护。也可以选择由 SSL 为数据提供保护。只有在初始手动验证过程后，设备才能复制到其他设备/从其他设备进行复制，请参见[“创建和编辑目标” \[330\]](#)。

复制包含下列已知限制：

- 更改目标 IP 地址会破坏复制
- 操作无法在池之间移动
- 每个项目级复制的 I/O 最大受限速度为 200 MB/s

了解复制

复制术语

- 复制对等设备（或此上下文中的对等设备）：已配置为复制源或目标的 ZFS Storage Appliance。
- 复制源（或源）：包含要复制到另一 ZFSSA 对等设备（目标）的数据的 ZFSSA 对等设备。每个 ZFSSA 既可以充当源也可以充当目标，但在一个特定复制操作的上下文中只能充当其中一种。
- 复制目标（或目标）：将接收并存储从另一 ZFSSA 对等设备（源）中复制的数据的 ZFSSA 对等设备。此术语也指 ZFSSA 上使得设备可以向另一 ZFSSA 进行复制的一个配置对象。
- 复制组（或组）：作为一个单元进行复制的数据集的集合（一个项目和一定数目的共享资源）。请参见[“项目级与共享资源级复制” \[329\]](#)。
- 复制操作（或操作）：源 ZFSSA 上的一个配置对象，其指定了项目或共享资源、目标 ZFSSA 和策略选项（包括发送更新的频率、是否对网络上的数据进行加密等等）。
- 数据包：某个操作的目标端对应体；目标 ZFSSA 上的一个配置对象，其管理作为某个特定操作的一部分从特定源复制的数据。源 ZFSSA 上的每个操作都恰好与目标 ZFSSA 上的一个数据包相关联，反之亦然。丢失了任何一个对象都将要求创建新的操作/数据包对（和完整的复制更新）。
- 完整同步（或完整更新）：发送项目及其某些共享资源的完整内容的一个复制操作。

- 增量更新：仅发送自上次更新以来项目及其共享资源中的差异的复制操作（无论上次更新是完整更新还是增量更新）。

项目复制目标

在可以将源 ZFSSA 复制到目标之前，这两个系统都必须设置一个复制对等设备连接，以使 ZFSSA 能够在将来的通信中安全识别对方。管理员可以通过在源 ZFSSA 的 "Configuration"（配置）> "Services"（服务）> "Remote Replication"（远程复制）屏幕上创建一个新的复制目标来设置此连接。要创建新目标，管理员需指定以下三个字段：

- 名称（仅用于在源 ZFSSA 的 BUI 和 CLI 中识别目标）
- 网络地址或主机名（用以联系目标 ZFSSA）
- 目标 ZFSSA 的 root 用户密码（用以授权管理员在目标 ZFSSA 上设置连接）

之后，ZFSSA 将交换用于在随后的通信中安全识别对方的密钥。这些密钥将作为 ZFSSA 配置的一部分永久存储，并且在重新引导和升级后也一直保留。如果 ZFSSA 恢复了出厂设置或得到重新安装，将丢失这些密钥。root 用户密码从不永久存储，因此在任一 ZFSSA 上更改 root 用户密码不需要对复制配置进行任何更改。因为此初始身份交换（例如所有复制控制操作）受 SSL 保护，所以密码从不以明文形式进行传输。

默认情况下，复制目标连接不是双向的。如果管理员配置了从源 A 到目标 B 的复制，则 B 无法自动将 A 作为目标。不过，系统支持反转复制方向，这将自动在 B 上为 A 创建一个目标（如果尚未存在），以便 B 可以复制回 A。

注意：如果复制源使用 NIS 或 LDAP 服务映射用户或用户群且此类用户或用户群包含在源上的共享资源配置中（例如在 "Share Level ACL"（共享资源级 ACL）或 "Share Space Usage"（共享资源空间利用率）中），此类用户或用户群应可作为复制目标（例如通过使用同样的 NIS 或 LDAP 服务器），否则复制服务器/反转操作会失败。

要配置复制目标，请参见[“配置项目复制” \[329\]](#)。

项目复制操作和数据包

目标在 ZFSSA 间提供了一个连接来使设备可以出于复制目的进行安全的通信，但是目标没有指定将复制的对象、频率或要使用的选项。因此，管理员必须在源 ZFSSA 上定义复制操作。操作是用于复制的主要管理控制点，每个操作将指定：

- 复制组（项目和一定数目的共享资源）
- 目标 ZFSSA
- 目标 ZFSSA 上的存储池（仅在初始设置期间使用）
- 频率（可以为手动、调度或连续）
- 其他选项，例如是否对网络上的数据流进行加密

组是按配置了操作的项目或共享资源隐式指定的（请参见[“项目级与共享资源级复制” \[329\]](#)）。创建操作后，无法更改目标 ZFSSA 和存储池，但您可以随时更改其他选项。通常，如果在更改某个选项时复制更新正在进行，则新值只有在下次更新开始时才会生效。

操作是 ZFSSA 上的主要复制配置单元。每个操作对应于目标 ZFSSA 上的一个数据包，该数据包中包含自上次复制更新开始以来配置了操作的源项目和共享资源的精确副本。管理员可以通过修改相应操作的属性来配置复制更新的频率和其他选项。在源 ZFSSA 上创建操作将会在目标 ZFSSA 上的指定的存储池中创建数据包，因此，在最初创建操作时，源必须能够与目标进行通信。

对每个复制操作进行首次更新时，会发送完整同步（或完整更新）：该操作的项目和共享资源的所有内容将被发送到目标 ZFSSA。此初始同步完成后，后续复制更新将以增量方式进行：仅发送自上次更新以来发生更改的内容。操作（在源上）和数据包（在目标上）通过指定的复制快照来跟踪已将哪些更改复制到目标中。通常，只要为操作发送了至少一个完整同步，并且操作/数据包连接未由于软件故障或管理操作而被破坏，复制更新就将以增量方式进行。

操作和数据包相互绑定。如果数据包因某种原因被破坏或销毁，那么即使该目标仍具有与此操作关联的数据和快照，此操作也无法发送复制更新。同样，如果操作被销毁，则数据包将无法接收新的复制更新（即使该源仍具有相同的数据和快照）。对于尝试执行会销毁“操作-数据包”连接的操作的管理人员，BUI 和 CLI 会对其发出警告。如果某个错误或显式管理操作破坏了“操作-数据包”连接，以致无法再进行增量更新，则管理员必须销毁数据包和操作或者断开其连接，并在源上创建一个新操作。

注意：除非管理员明确请求，否则 ZFSSA 会避免销毁目标上的数据。因此，如果某项操作的初始复制更新在复制某些数据后失败，在数据包中留下了不完整的数据，则使用同一操作的后续复制更新也将失败，因为 ZFSSA 无法覆写之前接收的数据。要解决这个问题，管理员应销毁现有的操作和数据包，并创建新的操作和数据包，然后再次开始复制。

在 2010.Q1 之前的软件发行版中，操作和副本配置（例如目标配置）存储在控制器上，而不是作为项目和共享资源配置的一部分存储在存储池中。因此，恢复出厂设置会导致配置被销毁。在 2010.Q1 及更高的发行版中，操作和数据包配置随对应的项目和共享资源存储在存储池中，因此即使在恢复出厂设置后，也仍然可用。不过，目标信息仍将丢失，并且目前无法将缺少目标的操作配置为指向新目标。

项目复制存储池

最初配置操作时，会让管理员选择目标上的哪个存储池应当包含复制的数据。创建某个操作后，包含该操作的存储池将无法更改。创建操作会在目标上在指定的存储池中创建一个空数据包，执行此操作之后，源不知道目标上的存储配置。它不跟踪此操作被复制到哪个池，也不使用目标上的存储配置更改进行更新。

当目标是一个属于群集的系统时，选择的存储池必须属于拥有用于复制的源使用的 IP 地址的同一个机头，因为在源使用该 IP 地址与目标进行联系时，只能始终确保可访问这

些池。这与 NAS 客户机 (NFS 和 SMB) 的配置完全相同, 其中, 在挂载操作中请求的 IP 地址和路径必须遵守相同的约束。在执行会更改群集中存储池和 IP 地址的所有权的操作时, 管理员必须考虑对复制到群集的源的影响。当前无法在存储池之间移动数据包。

项目级与共享资源级复制

ZFSSA 允许管理员在项目级别和共享资源级别上配置远程复制。与可在 "Shares" (共享资源) 屏幕上配置的其他属性一样, 每个共享资源均可以继承或覆写其父项目的配置。继承配置意味着共享资源不仅与父项目一样按相同的调度表使用相同的选项复制到相同的目标, 而且还将同一流中使用与继承了项目配置的其他共享资源相同的项目级快照进行复制。对于要求存储在多个共享资源上的数据之间保持一致性的应用程序, 这可能很重要。覆盖配置意味着不会使用任何项目级操作来复制此共享资源, 虽然可以使用其自身将包括项目的共享资源级操作来进行复制。不能覆盖项目的一部分复制配置, 然后继承其余配置。

更确切地说, 项目及其共享资源的复制配置定义了一定数目的复制组, 其中每个组都是通过单个流使用同时创建的快照复制的。所有组都包含项目本身 (实质上仅包括其属性)。一个项目级组包括继承了父项目的复制配置的所有共享资源。覆盖项目配置的所有共享资源构成了一个仅包含项目和共享资源本身的新组。

例如, 假定我们拥有以下内容:

- 项目 home 和共享资源 bill、cindi 以及 dave。
- home 的复制配置有一定数目的操作
- home/bill 和 home/cindi 继承项目复制配置
- home/dave 使用其自身具有一定数目的操作的配置覆盖了项目的复制配置

此配置定义了以下复制组, 其中每个组在每次操作时都作为单个流使用在项目和共享资源上同时创建的快照进行复制:

- 一个包括 home、home/bill 和 home/cindi 的项目级组。
- 一个包括 home 和 home/dave 的共享资源级组。

考虑到当前限制, 请勿在同一项目中混合进行项目级复制和共享资源级复制。这避免在反转复制方向或复制克隆时出现不可预知的结果。有关更多信息, 请参见[“管理复制数据包” \[335\]](#)和[“复制克隆” \[354\]](#)章节。

配置项目复制

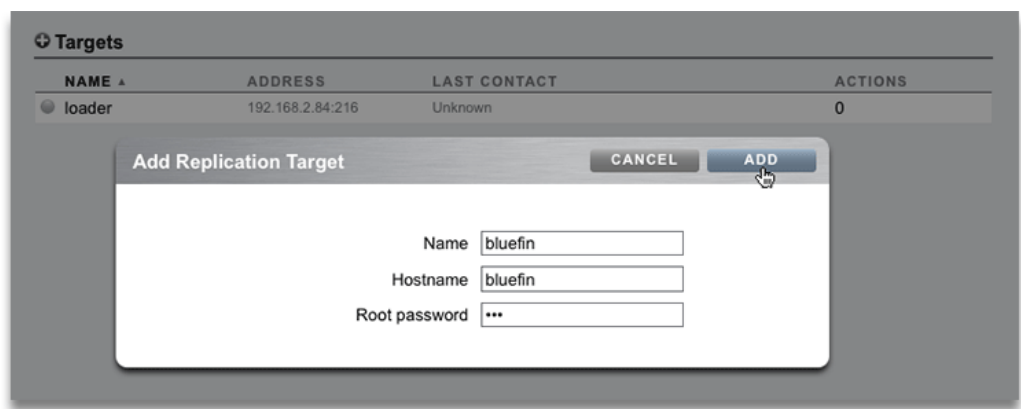
在配置复制之前, 请确保阅读并理解了有关复制目标、操作和数据包的上述各章节。

创建和编辑目标

该章节描述了创建和编辑目标。

▼ 在 BUI 中创建和编辑目标

- 1. 要在 BUI 中创建远程复制目标，请转至 "Configuration"（配置）> "Services"（服务）> "Remote Replication"（远程复制）> "Targets"（目标）。单击  "Targets"（目标）并配置 "Name"（名称）、"Hostname"（主机名）和 "Password"（密码）。
- 2. 要在 BUI 中编辑远程复制目标，请转至 "Configuration"（配置）> "Services"（服务）> "Remote Replication"（远程复制）> "Targets"（目标）。针对要编辑的目标，将光标移至目标名称上方，单击铅笔图标，然后配置 "Name"（名称）和/或 "Hostname"（主机名）。主机名必须解析为与之前相同的 ZFSSA（按目标的序列号核对）。如果要指向的 ZFSSA 不同于之前配置的 ZFSSA，则必须创建一个新目标以针对新 ZFSSA 进行验证。



▼ 在 CLI 中创建和编辑目标

- 1. 在 CLI 中，导航到 **targets** 节点，设置或取消设置目标 **hostname**、**root_password** 和 **label**。

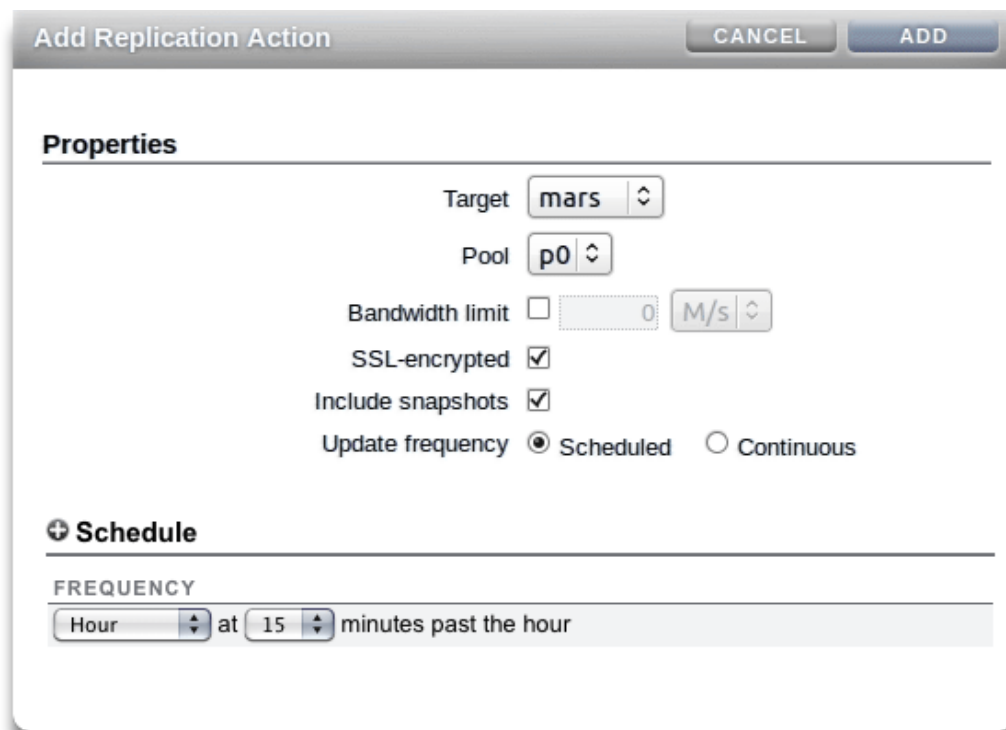
```
knife:> configuration services replication targets
```
- 2. 从该上下文中，管理员可以：

- 添加新目标
 - 查看随现有目标配置的操作
 - 编辑目标的唯一标识符（标签）和/或主机名
 - 销毁某个目标（如果没有操作在使用它）
3. 当有操作正在使用目标时，不应将其销毁。这将永久破坏此类操作。系统会尽最大努力强制实施此约束，但无法确保导出的存储池中不存在正使用给定目标的任何操作。

创建和编辑操作

复制操作具有以下属性，这些属性在 BUI 和 CLI 中以略微不同的形式提供：

图 13-1 添加复制操作



The image shows a graphical user interface window titled "Add Replication Action". At the top right are "CANCEL" and "ADD" buttons. The main content area is divided into two sections: "Properties" and "Schedule".

Properties

- Target:** A dropdown menu showing "mars".
- Pool:** A dropdown menu showing "p0".
- Bandwidth limit:** A checkbox that is unchecked, followed by a text input field containing "0" and a unit dropdown menu showing "M/s".
- SSL-encrypted:** A checkbox that is checked.
- Include snapshots:** A checkbox that is checked.
- Update frequency:** Two radio buttons: "Scheduled" (which is selected) and "Continuous".

+ Schedule

FREQUENCY

A dropdown menu showing "Hour" is followed by the text "at", then a text input field containing "15", followed by the text "minutes past the hour".

表 13-1 复制操作 CLI 属性

属性 (CLI 名称)	说明
Target (目标)	复制目标系统的唯一标识符。此属性是在最初配置操作时指定的，且此后将不再变动。
Pool (池)	目标上要将此项目复制到的存储池。此属性是在最初配置操作时指定的，且此后不再显示。
Mode (模式) (CLI : continuous) 和 schedule (调度表)	将以连续、手动或按调度的时间间隔复制此操作。有关详细信息，请参见 “复制模式：调度或连续” [335] 。
Include Snapshots (包含快照)	复制更新是否包括非复制快照。有关详细信息，请参见 “复制 – 包括中间快照” [335] 。
Limit bandwidth (限制带宽)	指定此复制更新的最大速度（按在网络上每秒传输的数据量计）。在复制更新期间对该属性所做的更改不会生效，直至下次更新。
Bytes sent (发送的字节数)	只读属性，描述发送到目标的字节数。
Estimated size (估计大小)	只读属性，描述估计要复制的数据大小。
Estimated time left (估计剩余时间)	只读属性，描述估计完成前的剩余时间。
Average throughput (平均吞吐量)	只读属性，描述平均复制吞吐量。
Use SSL (使用 SSL)	是否使用 SSL 对网络上的数据进行加密。使用此功能会对每个操作的复制性能产生显著影响。
State (状态)	只读属性，描述操作当前是否处于空闲、正在发送更新或正在取消更新的状态。
Last sync (上次同步)	只读属性，描述上次成功发送更新的时间。如果自引导后系统尚未发送成功更新，则该值可能为 "unknown"。
Last attempt (上次尝试)	只读属性，描述上次尝试更新的时间。如果自引导后系统尚未尝试发送更新，则该值可能为 "unknown"。
Next update (下次更新)	只读属性，描述下次进行尝试的时间。此值可以是一个日期（对于调度的更新）、"manual"（手动）或 "continuous"（连续）。

▼ 在 BUI 中创建和编辑操作

1. 在至少配置了一个复制目标后，通过在 BUI 中导航到本地项目或共享资源并单击 "Replication"（复制）选项卡，或通过在 CLI 中导航到本地项目或共享资源并选择 "replication" 节点，管理员可以在该本地项目或共享资源上配置操作。这些接口显示在项目或共享资源上配置的现有操作的状态和复制进度信息，便于管理员创建新操作：

Actions		
TARGET ▲	UPDATES	STATUS
jupiter Manual	2013-7-15 17:22:04 Synced 2013-7-15 17:22:04 Attempted	Sync now
jupiter Continuous	2013-7-15 17:24:48 Synced 2013-7-15 17:25:21 Fail	86% of 3.5G @21MB/s (-00:00:24)
mars Scheduled	2013-7-15 17:23:32 Synced 2013-7-15 17:23:32 Attempted	2013-7-15 17:34:00 Next
venus Manual	2013-7-15 17:18:43 Synced 2013-7-15 17:18:43 Attempted	74% of 4.0G @24MB/s (-00:00:45)

- 复制到目标时，将显示两行状态信息。第一行显示目标名称、上次成功同步的日期和时间以及进度指示条，如果是连续复制，则显示螺旋条状的进度指示条。第二行显示复制类型（调度、手动或连续）、上次尝试或失败同步的日期和时间以及状态详细信息。对于正在进行的复制，状态详细信息包含完成百分比、估计要复制的数据大小、平均复制吞吐量和估计完成时间。复制未在进行时，“Status”（状态）列将根据复制类型显示下次调度的复制或“Sync now”（立即同步）消息。

▼ 在 CLI 中创建和编辑操作

- 在 CLI 中可以显示相同的进度信息，对于正在进行的复制，将显示状态 **sending**：

```
otoro:shares otoro-proj-01 action-000> show
Properties:
    id = 80a96f4f-93fe-4abd-eb54-fb82e7f8c69f
    target = chutoro
    continuous = false
    include_snaps = true
    max_bandwidth = unlimited
    bytes_sent = 505M
    estimated_size = 3.0G
    estimated_time_left = 00:00:41
    average_throughput = 63MB/s
    use_ssl = false
    state = sending
    state_description = Sending update
    next_update = Sync now
    last_sync = Sun Jul 14 2013 06:04:38 GMT+0000 (UTC)
    last_try = Sun Jul 14 2013 06:04:38 GMT+0000 (UTC)
    last_result = success
```

2. 注意：根据复制的数据大小，复制可能需要很长时间才能完成。使用进度信息可以确定更新状态。对于初始复制，一定不要将其中断，包括但不限于重新启动 ZFSSA 或取消更新；否则，必须重新启动整个初始复制。
3. 复制目标信息可显示在 CLI 中，状态为 **actions**：

```

otoro:configuration services replication targets> show

Targets:
  TARGET      LABEL      ACTIONS
  target-000  oakmeal    1

otoro:configuration services replication targets> select target-000

otoro:configuration services replication target-000> show
Properties:
  address = 10.153.34.167:216
  label = oakmeal
  hostname = oakmeal-7320-167
  asn = 4913649f-7549-6d2a-866b-987ddbc4e163
  actions = 1

oakmeal-7320-167:configuration services replication target-000> actions
  POOL      PROJECT      SHARE
  pool1     project1     (multiple)

```

4. 使用 CLI 时，知道新创建的复制操作的 ID 很有帮助。后来，可使用 ID 选择正确的复制操作节点。要查看新创建的操作的 ID，请使用 **last** 命令导航到新复制操作的节点。然后使用 **get id** 命令检索操作 ID：

```

otoro:> shares
otoro:shares> select p1
otoro:shares p1> replication
otoro:shares p1 replication> create
otoro:shares p1 action (uncommitted)> set target=oakmeal
  target = oakmeal (uncommitted)
otoro:shares p1 action (uncommitted)> set pool=p
  pool = p (uncommitted)
otoro:shares p1 action (uncommitted)> set use_ssl=false
  use_ssl = false (uncommitted)
otoro:shares p1 action (uncommitted)> commit
otoro:shares p1 replication> last
otoro:shares p1 action-001> get id
  id = fb1bb3fd-3361-42e1-e4a1-b06c426172fb
otoro:shares p1 action-001> done
otoro:shares p1 replication>

```

复制模式：调度或连续

可将复制操作配置为按调度表或连续发送更新。在两种情况下，复制更新过程本身均相同。此属性仅控制时间间隔。

由于连续复制操作以尽可能频繁的方式发送更新，因此它们会导致向目标系统发送包含所有文件系统更改的恒定流。对于变动较大的文件系统（在较短的时间间隔内创建并销毁大量文件），这会导致复制的数据远远超过实际需要复制的数据。不过，只要复制与数据更改保持一致，此模式可保证在源系统发生数据丢失灾难时丢失的数据最少。

连续复制仍然是异步的。ZFS Storage Appliance 当前不支持同步复制，同步复制在数据已同时提交到主存储系统和辅助存储系统上的稳定存储之前，不考虑提交到稳定存储的数据。

复制 – 包括中间快照

当 "Include Snapshots"（包含快照）属性为 true 时，复制更新包括自上次复制更新后（如果是首次完整更新，则是自创建共享资源后）创建的非复制快照。其中包括自动快照和管理员创建的快照。可以禁用此属性，以跳过这些快照，并且在每次更新时仅发送复制快照之间的更改。

复制 – 发送和取消更新

对于配置有调度的复制或手动复制的目标，通过在 BUI 中单击  按钮或在 CLI 中使用 sendupdate 命令，管理员可以立即发送复制更新。如果正在发送某个更新，则此功能不可用（或不起作用）。在发送更新之前，请确保目标上有足够的磁盘空间来复制整个项目。

如果更新当前处于主动状态，则 BUI 将显示一个进度指示条，CLI 将显示 sending 状态。要取消更新，请单击  按钮或使用 cancelupdate 命令。完成取消可能需要花费几秒钟的时间。

管理复制数据包

数据包是复制的项目和共享资源的容器。源 ZFSSA 上的每个复制操作都与目标 ZFSSA 上的一个数据包相对应，如上文所述。通过 BUI 和 CLI，管理员都可以像浏览本地项目和共享资源一样浏览复制的项目、共享资源、快照和属性。不过，由于复制的共享资源必须与源 ZFSSA 上的对应共享资源完全匹配，因此在复制数据包中许多管理操作都不

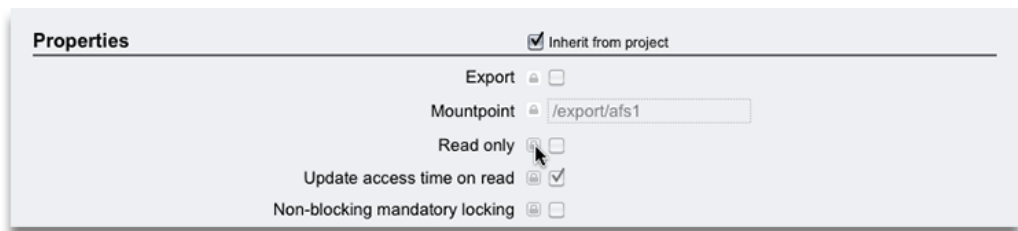
允许执行，这包括创建、重命名和销毁项目和共享资源，创建和重命名快照，以及修改项目和共享资源的大多数属性。在复制数据包中，可以销毁未用作增量复制的基础的快照。通常不建议采用此做法，但在需要额外的可用空间时可以使用。

在 2009.Q3 及更早的软件版本中，无法在已复制的共享资源中更改属性。2010.Q1 发行版（与关联的延迟升级）针对修改已复制的共享资源的属性增加了有限支持，以便在源和目标 ZFSSA 上实施不同的策略。此类属性修改在每次复制升级后也一直保留。只能修改已复制的项目和共享资源的以下属性：

- 预留、压缩、副本、重复数据删除和缓存。可以在复制目标上更改这些属性，以影响目标 ZFSSA 上来自源的各种成本、灵活性、性能或可靠性策略。
- 挂载点和共享属性（例如 sharenfs、SMB 资源名称等）。这些属性控制如何将共享资源导出到 NAS 客户机以及如何更改共享资源以在来自源的目标 ZFSSA 上影响不同的安全或保护策略。
- 自动快照策略。可以在目标系统上更改自动快照策略，但这些更改在断开数据包连接之前不会产生任何影响。在已复制的项目和共享资源上不会生成或销毁自动快照。

BUI 和 CLI 不允许管理员更改不可变的属性。对于共享资源，使用了一个不同的图标来指示无法更改属性继承：

图 13-2 管理复制数据包属性



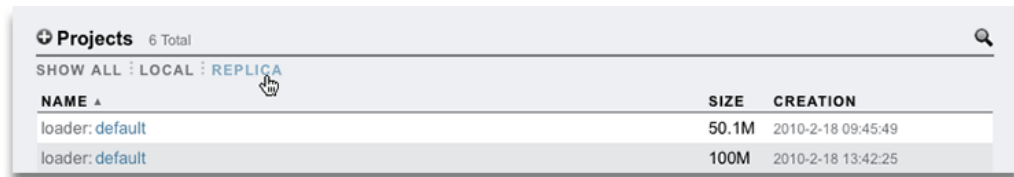
要在这类目标上修改属性，必须在复制目标上应用随 2010.Q1 发行版提供的延迟更新。系统将不允许管理员在未应用 2010.Q1 延迟更新的系统上修改复制数据包内的属性。

当前发行版不支持“链式”复制的配置（也就是说，将已复制的共享资源复制到其他 ZFSSA）。

管理 BUI 中的复制数据包

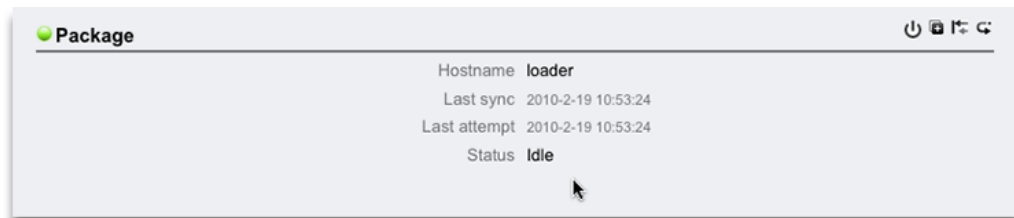
复制数据包作为项目显示在 BUI 中的 "Replica"（副本）过滤器下：

图 13-3 "Replica" (副本) 过滤器



选择要编辑的复制数据包后，管理员将进入数据包的项目的 "Shares" (共享资源) 视图。由此，管理员可以像管理本地共享资源那样管理已复制共享资源（上述例外除外）。在 "Replication" (复制) 选项卡下，可以修改数据包属性（包括状态）：

图 13-4 数据包项目的 "Shares" (共享资源) 视图



复制失败后，左侧的状态图标会发生变化：

图 13-5 状态图标显示失败



在首次复制更新开始后，数据包只显示在 BUI 中。在首次更新完成后的某个时刻之前，这些数据包可能不会显示在列表中。

在 CLI 中管理复制数据包

在 CLI 中，复制数据包按源组织在 `shares replication sources` 下方。管理员先选择一个源，然后选择一个数据包。在此节点上可以执行数据包级操作，还可以选择项目以像管理本地项目和共享资源那样管理项目属性和共享资源（上述为例外情况）。例如：

```
loader:> shares replication sources
loader:shares replication sources> show
Sources:

source-000 ayu
      PROJECT   STATE      LAST UPDATE
package-000 oldproj   idle       unknown
package-001 aproj1    receiving  Sun Feb 21 2010 22:04:35 GMT+0000 (UTC)

loader:shares replication sources> select source-000
loader:shares replication source-000> select package-001
loader:shares replication source-000 package-001> show
Properties:
      enabled = true
      state = receiving
      state_description = Receiving update
      last_sync = Sun Feb 21 2010 22:04:40 GMT+0000 (UTC)
      last_try = Sun Feb 21 2010 22:04:40 GMT+0000 (UTC)

Projects:
      aproj1

loader:shares replication source-000 package-001> select aproj1
loader:shares replication source-000 package-001 aproj1> get mountpoint
      mountpoint = /export
loader:shares replication source-000 package-001 aproj1> get sharenfs
      sharenfs = on
```

您也可通过 `configuration services replication` 显示复制源。例如：

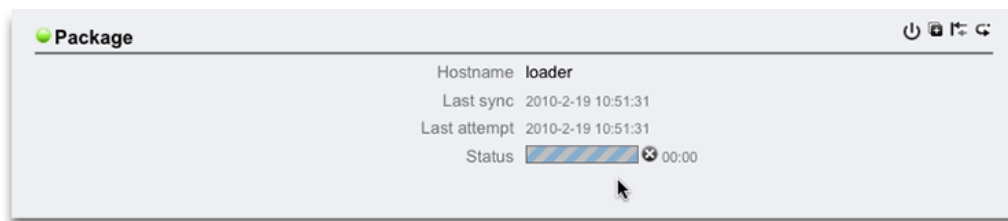
```
loader:configuration services replication> show
Properties:
      <status> = online

Children:
      targets => Configure replication targets
      sources => View and manage replication packages
```

取消复制更新

要使用 BUI 取消目标中的正在进行的复制更新，请首先导航到复制数据包（参见上文），然后单击 "Replication"（复制）选项卡。如果某个更新正在进行，则如下所示，您将看到在取消按钮 (✖) 旁边有一个螺旋条状的进度条：

图 13-6 取消复制



单击此按钮可取消更新。

要使用 CLI 取消目标中正在进行的复制更新，请首先导航到复制数据包（参见上文），然后使用 `cancelupdate` 命令。

可能无法从目标上启动更新。管理员必须登录到源系统才能启动手动更新。

禁用数据包

可以完全禁用某个数据包的复制更新，此操作将取消任何正在进行的更新并导致来自源 ZFSSA 的新更新失败。

要从 BUI 切换数据包的启用/禁用状态，请首先导航到数据包（参见上文），单击

"Replication"（复制）选项卡，然后单击  图标。左侧的状态图标应当会变化以指示数据包的状态（已启用、已禁用或失败）。数据包将一直保持禁用状态，直到由管理员使用同一按钮或 CLI 显式启用。

要从 CLI 切换数据包的启用/禁用状态，请首先导航到数据包（参见上文），修改 `enabled` 属性，然后提交更改。

克隆数据包或各个共享资源

已复制的数据包的克隆是一个本地的可变项目，可以像系统上的任何其他项目一样进行管理。克隆的共享资源是在最新接收的快照中复制的共享资源的克隆。这些克隆以与共享资源快照的克隆相同的方式与其原始快照一起共享存储（请参见“[克隆快照](#)” [303]）。此机制可用于在复制源中出现灾难性问题时进行故障转移，或者只是提供可以修改的数据本地版本。

可以使用 BUI 中的  按钮或使用 `clone` CLI 命令（在数据包的上下文中）基于最新接收的复制快照创建数据包克隆。CLI 和 BUI 接口都要求管理员指定新克隆项目的名称，

并且允许管理员覆盖项目或其共享资源的挂载点，以确保这两者与系统中的其他共享资源的挂载点不冲突。

在 2009.Q3 及更早版本中，克隆已复制的项目是访问其数据的唯一方法，因此这也是实施灾难恢复故障转移的唯一方法。在 2010.Q1 及更高版本中，可以将各个文件系统以只读形式导出而无需创建克隆。此外，作为故障转移操作的一部分，可以将复制数据包直接转换为可写的本地项目。因此，不再需要克隆数据包也不再建议您这样做，因为这些替代方法使用更简单的操作提供了类似功能，且无需管理克隆及其依赖项。

特别是，当存在克隆时，无法销毁其原始快照。销毁快照（可能是由于销毁快照所属的共享资源、项目或复制数据包而导致的结果）时，系统会向管理员发出警告以指明该操作将销毁任何依赖它的克隆。请注意，作为后续复制更新的一部分，还可以随时在源上销毁快照，以及在目标上销毁此类快照。如果此类快照拥有克隆，则快照将使用唯一名称进行重命名（通常为 recv-XXX）。

管理员还可以使用常规的 BUI 和 CLI 接口克隆单个已复制的共享资源快照。

导出已复制的文件系统

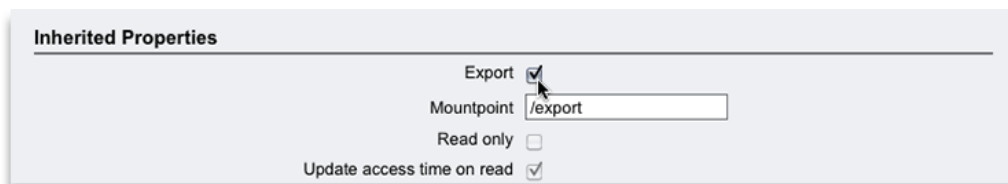
可以将已复制的文件系统以只读形式导出到 NAS 客户机。这可用于验证已复制的数据或对已复制的数据执行备份或其他密集型操作（转移源 ZFSSA 的此类工作负荷）。

文件系统的内容始终与最新接收的该文件系统的复制快照相匹配。这可能比最新接收的整个数据包的快照要新，且它可能与在同一个数据包中的其他共享资源的最新快照不匹配。有关详细信息，请参见“[快照和数据一致性](#)” [352]。

复制更新是在文件系统级别自动应用的。查看复制的文件的客户机会将复制更新视为底层文件系统上的瞬时更改。使用在最新更新中删除了的文件进行工作的客户机将收到错误。使用在最新更新中更改了的文件进行工作的客户机将立即看到更新的内容。

默认情况下，不会导出已复制的文件系统。可以通过使用 BUI 或 CLI 修改项目或共享资源的 "exported"（已导出）属性来将其导出。

图 13-7 继承的属性



此属性像其他共享资源属性一样进行继承。不会为本地项目和共享资源显示此属性，因为它们始终被导出。此外，断开复制连接（这会将数据包转换为本地项目）将导致数据包的共享资源变为已导出状态。

当前无法导出已复制的 LUN。必须首先克隆已复制的 LUN 或者断开复制数据包的连接，然后才能导出其内容。

断开复制连接

通过断开复制连接，可以将复制数据包转换为其行为方式与其他本地项目完全相同的一个本地可写项目（即，没有应用于复制数据包的管理限制）。执行此操作后，由于此数据包将无法再接收复制更新，因此来自源中的同一个项目的后续复制更新将需要执行新操作来发送完整更新（发送到新的数据包）。使用同一操作的后续复制更新将失败，因为对应的数据包不再存在于目标中。

当使用复制在 ZFSSA 之间迁移数据时或者在不涉及将接收到的数据复制回源（作为典型的双系统灾难恢复计划的一部分）的其他情形中，此选项非常有用。

在 BUI 中可通过以下方法来断开复制连接，方法如下：导航到复制数据包（请参见上文），单击 "Replication"（复制）选项卡，然后单击  按钮。通过显示的对话框，管理员可以指定新本地项目的名称。

在 CLI 中可通过以下方法来断开复制连接：导航到复制数据包（参见上文），然后使用 `sever` 命令。此命令接受一个用于指定新本地项目名称的可选参数。如果未指定参数，则将使用原始名称。

因为所有的本地共享资源都将导出，所以在断开数据包连接时将导出数据包中的所有共享资源，无论之前是否导出了这些共享资源（请参见上文）。如果复制的文件系统与系统上的其他文件系统之间存在挂载点冲突，则该断开连接操作将失败。在断开连接之前，必须通过重新配置相关共享资源的挂载点来解决这些冲突。

反转复制方向

可以反转复制方向以支持典型的双系统灾难恢复计划。此操作与上文介绍的断开连接操作类似，但需要额外在新的本地项目上配置复制操作，以便向源系统进行增量复制。完成此操作时不会在源系统上进行任何更改，但使用此操作的首次更新尝试会将源系统上的原始项目转换为一个复制数据包，并回滚自上次成功从该系统进行复制更新后所做的任何更改。

除了修改主数据副本和辅助数据副本的读写状态以外，该功能不会自动重定向生产工作负荷、对 IP 地址进行故障转移，或执行与灾难恢复故障转移相关的其他活动。

在将原始源项目转换为原始源系统（现在充当目标）上的复制数据包的过程中，作为当前正在反转的操作/数据包的一部分复制的共享资源将移动到新的复制数据包中，且不会

导出。原始项目仍保留在本地集合中，但如果操作/数据包包含其所有的共享资源，则它最后可能会成为空的。反转共享资源级复制时，原始项目中的所有其他共享资源保持不变。

在建立了从一台 ZFSSA 到另一台 ZFSSA 的共享资源级复制后，反转目标设备上的复制会销毁复制调度表。然后，会在项目级别上创建一个复制操作，它包含正确的目标 ZFSSA，但无调度表。

如上文所述，该功能通常用于实施双系统灾难恢复配置，在这种配置中，主系统提供生产数据并将其复制到辅助或 *DR* 系统（通常位于另一数据中心），而后者准备在主站点发生灾难时接管生产通信。当主站点发生灾难时，必须将辅助站点的副本设为“主”副本，方法如下：将它设置为可写的，并将生产通信重定向到辅助站点。当主站点修复后，可以将辅助站点上累积的更改复制回主站点，且该站点可继续为生产工作负荷提供服务。

此类计划中的典型事件序列如下所述：

- 主系统正在为生产工作负荷提供服务，且正在向辅助系统进行复制。
- 发生了一个灾难，可能表示在主站点上发生了总系统故障。管理员在辅助站点上反转复制方向，并且将复制的共享资源导出到为了在主服务恢复后将数据复制回主站点而配置的一个新项目中。与此同时，生产工作负荷被重定向到辅助站点。
- 主站点重新联机后，管理员启动从辅助站点到主站点的复制更新。这会将主站点的副本转换为一个复制数据包，回滚自上次成功更新目标后（发生故障前）所做的任何更改。获得最新的主站点副本时，管理员将再次反转复制方向，使主站点上的副本具有可写性。生产通信被重定向到主站点。复制将恢复为从主站点到辅助站点，这就恢复了主副本和辅助副本之间的初始关系。

在反转数据包的复制方向时，强烈建议管理员先停止该项目从源的复制。如果在管理员反转项目的复制方向时复制更新正在进行，则管理员将无法得知在以前的目标 ZFSSA（现在是源 ZFSSA）上创建结果项目时使用的是哪个一致的复制快照。

在 BUI 中可通过以下方法来反转复制：导航到复制数据包（请参见上文），单击 "Replication"（复制）选项卡，然后单击  按钮。通过显示的对话框，管理员可以指定新本地项目的名称。

在 CLI 中可通过以下方法来反转复制：导航到复制数据包（参见上文），然后使用 `reverse` 命令。此命令接受一个用于指定新本地项目名称的可选参数。如果未指定参数，则将使用原始名称。

因为所有的本地共享资源都将导出，所以在反转数据包时将导出数据包中的所有共享资源，无论之前是否导出了这些共享资源（请参见上文）。如果复制的文件系统与系统上的其他文件系统之间存在挂载点冲突，则该反转操作将失败。在断开连接之前，必须通过重新配置相关共享资源的挂载点来解决这些冲突。由于此操作通常是恢复生产服务的关键路径的一部分，因此强烈建议您在首次设置系统时（而不是进行 DR 故障转移时）解决这些挂载点冲突。

销毁复制数据包

无法在未销毁整个数据包的情况下销毁数据包内的项目和共享资源。可以从 BUI 中通过销毁对应的项目来销毁整个数据包。可在 shares replication sources 节点上使用 destroy 命令从 CLI 中销毁数据包。

数据包被销毁后，来自相应操作的后续复制更新将失败。要恢复制，需要在源上重新创建操作，以便在要接收新数据副本的目标上创建新数据包。

复制任务

以下任务是复制过程的示例。

反转复制 – 建立复制

下面是反转复制支持典型的双系统灾难恢复的示例。在本例中，M11 是生产系统，M5 是恢复系统。

▼ 反转复制

1. 在生产系统 M11 上，导航到 "Configuration" (配置) > "SERVICES" (服务)。
2. 在 "Data Services" (数据服务) 下的 SMB 行，如果状态为 "Disabled" (已禁用)，请单击 "Enable service" (启用服务)。
3. 导航到 "Configuration" (配置) > "SERVICES" (服务) > "Remote Replication" (远程复制)。
4. 单击  "Targets" (目标)，然后配置名称、主机名和密码设置。Name=M5、Host name=192.168.1.17、Root password=pppp\$1234
5. 选择 Pool=Pool1。
6. 导航到 "Shares" (共享资源) > "PROJECTS" (项目)。
7. 单击  "Projects" (项目)。Name=P1
8. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Protocols" (协议)。

9. 在 SMB 章节，设置 Resource Name=on。
10. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Shares" (共享资源)。
11. 单击  "FileSystems" (文件系统)。Name=S1、User=root、Group=other、Permissions=RWX RWX RWX
12. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Shares" (共享资源) > "S1" > "Protocols" (协议)。SMB 章节显示可以使用 SMB 通过 \\192.168.1.7\S1 路径到达 S1。
13. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Replication" (复制)。
14. 单击  "Actions" (操作)，然后设置目标和池。Target=M5、Pool=Pool1
15. 单击  "Schedule" (调度表)，然后设置频率。Frequency=Half-Hour at 00 minutes past the hour.
16. 在 SMB 客户机系统上，映射网络驱动器 \\192.168.1.7\S1 (user=root, password=pppp \$1234)。
17. 创建名为 F1.txt 的文件。
18. 在生产系统 M11 上，导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Replication" (复制)。
19. 在 "Action" (操作) 行 TARGET=M5 上，单击 "Update now" (立即更新)。
20. 完成复制后单击 "Disable" (禁用)，"STATUS" (状态) 更改为已禁用。

反转复制 – 模拟从灾难中恢复

要模拟从阻止访问生产系统的灾难中恢复，请使用恢复系统反转复制。反转复制时，目标上存在的复制数据包被转换为一个本地项目，并且另外还为该本地项目配置了一个复制操作，用以向原始源系统进行增量复制。默认情况下不会启用该复制操作。管理员应手动发送更新。

▼ 反转复制

1. 要在 SMB 客户机系统上模拟断开与主系统 M11 的通信，请选择 "Disconnect network drive"（断开网络驱动器连接）。
2. 在灾难恢复系统 M5 上，选择 "Pool=Pool1"。
3. 导航到 "Shares" > "PROJECTS" > "REPLICA"。项目 M11:P1 已列出
4. 导航到 "Shares" > "PROJECTS" > "REPLICA" > M11:P1 > "Replication"。数据包的 "Status"（状态）为 "Idle"（闲置）
5. 单击 "Reverse the direction of replication"（反转复制方向），然后设置新的项目名称。New Project Name=P1
6. 导航到 "Shares" > "PROJECTS" > "REPLICA"。项目 M11:P1 不再列出，因为目标上存在的复制数据包已被转换为一个本地项目。
7. 导航到 "Configuration"（配置）> "Services"（服务）。
8. 在 "Data Services"（数据服务）下的 SMB 行，如果状态为 "Disabled"（已禁用），请单击 "Enable service"（启用服务）。
9. 导航到 "Shares"（共享资源）> "PROJECTS"（项目）> "LOCAL"（本地）。项目 P1 已列出。
10. 导航到 "Shares"（共享资源）> "PROJECTS"（项目）> "P1" > "Protocols"（协议）。
11. 在 SMB 章节，设置 Resource Name=on。
12. 导航到 "Shares"（共享资源）> "PROJECTS"（项目）> "P1" > "Shares"（共享资源）> "S1" > "Protocols"（协议）。SMB 章节显示可以使用 SMB 通过 \\192.168.1.17\S1 路径到达 S1。
13. 在 SMB 客户机系统上，映射网络驱动器 \\192.168.1.17\S1 (user=root, password=pppp \$1234)。
14. 编辑文件 F1.txt，并将其保存为 F2.txt。注意：在实际灾难恢复过程中，恢复与生产系统 M11 的通信后，将有机会在应用程序继续访问灾难恢复系统 M5 上数据的同时，触发手动、调度或连续复制更新。
15. 要为转回到生产系统做好准备，请选择 "Disconnect network drive"（断开网络驱动器连接）。

16. 在灾难恢复系统 M5 上，导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Replication" (复制)。
17. 在 "Action" (操作) 行 TARGET=M11 上，单击 "Update now" (立即更新)。
18. 完成复制后，单击 "Disable" (禁用)。

反转复制 – 从生产系统恢复复制

每次反转复制时，都要提供复制数据包转换为新本地项目时使用的新项目名称。如果要使用相同的名称，并且之前的复制反转留下一个空本地项目使用该名称，则必须删除现有空项目，以使下次反转可以创建同名的项目。

▼ 反转复制

1. 在生产系统 M11 上，导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "LOCAL" (本地) > "P1"。P1 为空，因为从新源系统 (原始目标) 进行的首次更新将原始源系统上的原始项目转换为一个复制数据包。本例利用项目级复制操作，复制项目中的所有共享资源。因此，原来位于该本地项目下的所有共享资源现在位于复制数据包下，使得本地项目为空。
2. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "LOCAL" (本地)。
3. 删除 P1。可以安全地删除该空项目，因为其内容在从原始目标进行反转复制后已随之移动到复制数据包中。
4. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "REPLICA" (副本) > M5:P1 > "Replication" (复制)。
5. 单击 "Reverse the direction of replication" (反转复制方向)，然后设置项目名称。New Project Name=P1
6. 在 SMB 客户机系统上，映射网络驱动器 \\192.168.1.7\S1 (user=root, password=pppp \$1234)。文件 F1.txt 和 F2.txt 已列出。
7. 在生产系统 M11 上，导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "P1" > "Replication" (复制)。
8. 在 "Action" (操作) 行 TARGET=M5 上，单击 "Edit entry" (编辑条目)。
9. 单击  "Schedule" (调度表)，然后设置频率。Frequency=Half-Hour at 00 minutes past the hour.

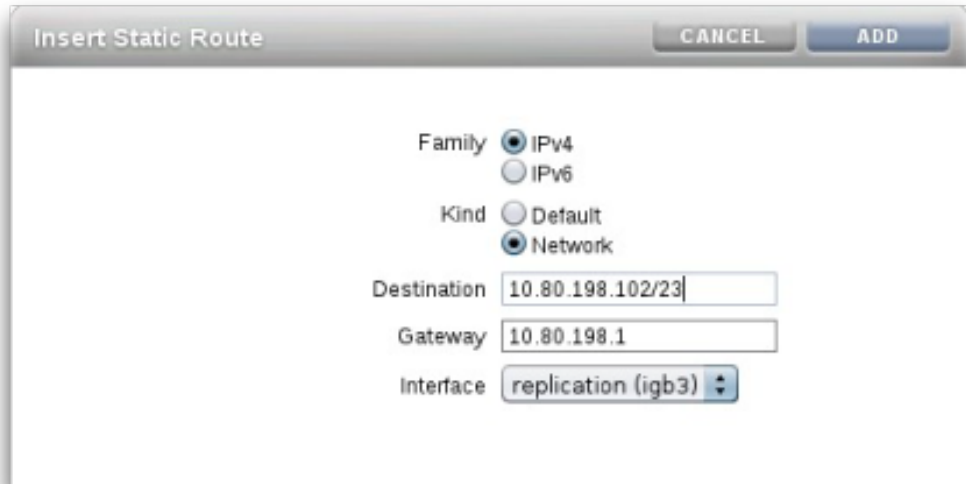
10. 在 "Action" (操作) 行 TARGET=M5 上, 单击 "Update now" (立即更新)。这样会指引灾难恢复系统 M5 将其本地项目 P1 转换回复制数据包 M11:P1。
11. 监视操作行 "TARGET=M5" 中的 "UPDATES" 列, 等待复制更新完成。
12. 在灾难恢复系统 M5 上, 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "LOCAL" (本地) > "P1"。P1 为空, 因为项目已转换回复制数据包。
13. 导航到 "Shares" (共享资源) > "PROJECTS" (项目) > "LOCAL" (本地)。
14. 要使下次反转将复制数据包转换为名为 P1 的项目, 请删除 P1。

强制复制以使用静态路由

要将复制通信合并到特定网络接口上, 您必须使用静态路由连接源和目标 ZFS ZFSSA。要设置静态路由, 请遵循以下步骤:

▼ 强制复制以使用静态路由

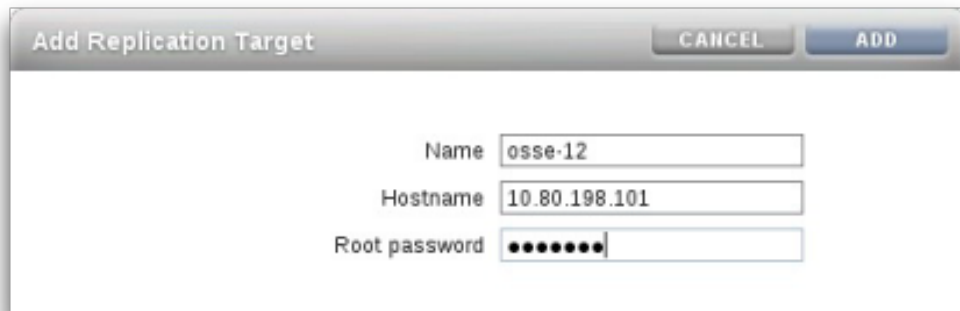
1. 要设置静态路由, 请在 "Configuration" (配置) > "Network" (网络) > "Routing" (路由) 页面单击  添加图标。
2. 在 "Insert Static Route" (插入静态路由) 方框中, 选择 "Family" (系列) 和 "Kind" (种类), 然后输入目标 IP、网关和接口。



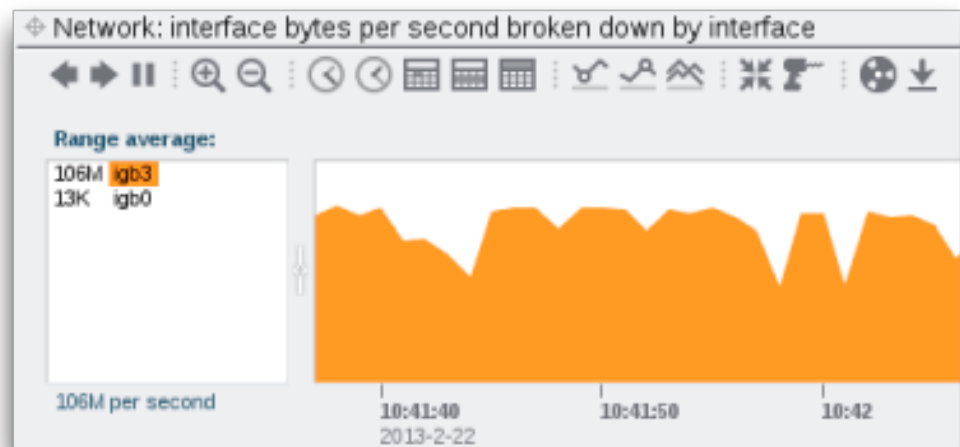
3. 单击 "Add" (添加) 。
4. 要确保通信经过源和目标，请在 CLI 中使用 `traceroute`。有关使用 `traceroute` 的信息，请参见[“网络路由配置” \[75\]](#)。在示例中，`10.80.219.124 @ igb0` 将 `igb0` 识别为接口。这是验证是否使用了正确接口的最快方式。

```
brmv01sn02:> traceroute poc7330-050
traceroute: Warning: Multiple interfaces found; using 10.80.219.124 @ igb0 traceroute
to poc7330-050 (10.80.219.117), 30 hops max, 40 byte packets
1  poc7330-050.us.oracle.com (10.80.219.117)  0.446 ms  0.115 ms  0.104 ms
```

5. 要添加新的复制目标，请在 "Configuration" (配置) > "Services" (服务) > "Replication" (复制) 页面单击  添加图标。
6. 在 "Add Replication Target" (添加复制目标) 方框中，键入目标名称、网络接口的主机名 IP 地址以及密码。



7. 单击 "Add" (添加) 。
8. 要确保通信经过定义的静态路由，请在复制开始后使用《[Oracle ZFS Storage Appliance Analytics 指南](#)》中的“[Network Interface Bytes \(网络 - 接口字节数\)](#)”。



9. "Preferences" (首选项) 页确保启用了“[首选项属性](#)” [137]。
10. 确认源至目标复制使用了正确的接口后，请反转复制。有关反转复制的信息，请参见“[反转复制方向](#)” [341]。

克隆接收的复制项目

下面是对接收的复制项目进行克隆（覆盖项目的挂载点和一个共享资源的挂载点）的一个 CLI 示例：

```
perch:> shares
perch:shares> replication
perch:shares replication> sources
perch:shares replication sources> select source-000
perch:shares replication source-000> select package-000
perch:shares replication source-000 package-000> clone
perch:shares replication source-000 package-000 clone> set target_project=my_clone
      target_project = my_clone
perch:shares replication source-000 package-000 clone> list
CLONE PARAMETERS
      target_project = my_clone
      original_mountpoint = /export
      override_mountpoint = false
      mountpoint =

      SHARE                                MOUNTPOINT
      bob                                (inherited)
      myfs1                              (inherited)
perch:shares replication source-000 package-000 clone> set override_mountpoint=true
      override_mountpoint = true
perch:shares replication source-000 package-000 clone> set mountpoint=/export/my_clone
      mountpoint = /export/my_clone
perch:shares replication source-000 package-000 clone> select bob
perch:shares replication source-000 package-000 clone bob> set override_mountpoint=true
      override_mountpoint = true
perch:shares replication source-000 package-000 clone bob> set mountpoint=/export/bob
      mountpoint = /export/bob
perch:shares replication source-000 package-000 clone bob> done
perch:shares replication source-000 package-000 clone> commit
CLONE PARAMETERS
      target_project = my_clone
      original_mountpoint = /export
      override_mountpoint = true
      mountpoint = /export/my_clone

      SHARE                                MOUNTPOINT
      bob                                /export/bob (overridden)
      myfs1                              (inherited)
Are you sure you want to clone this project?
There are no conflicts.
perch:shares replication source-000 package-000 clone>
```

远程复制详细信息

授权

除了在 "Services"（服务）范围下有一个允许管理员用来停止、启动和重新启动复制服务的 "Remote Replication"（远程复制）过滤器之外，复制子系统还在 Projects and Shares（项目和共享资源）范围下提供了两项“[用户授权](#)” [128]：

授权	详细信息
rrsource	允许管理员创建、编辑和销毁复制目标与操作，以及为复制操作发送和取消更新。
rrtarget	允许管理员管理复制的数据包，包括禁用数据包级复制、克隆数据包或其成员、修改接收到的数据集以及断开复制连接或反转复制。这些操作中的某些操作（例如设置属性或克隆各个共享资源）可能需要其他授权。有关详细信息，请参见 "Projects and Shares"（项目和共享资源）范围中的可用授权。

在 ZFSSA 上配置复制目标需要 rrsorce 授权，虽然这是在 "Remote Replication"（远程复制）服务屏幕下配置的。有关授权的帮助，请参见“[用户授权](#)” [128]。

警报

当发生以下任何事件时，系统将发布警报：

- 成功启动或完成了手动的或调度的复制更新（源和目标）。
- 任何复制更新失败时，包括因管理员显式取消而导致的复制更新失败（源和目标）。
- 因为同一操作的另一个更新已在进行而跳过了某个调度的复制更新（请参见上文）。
- 连续复制首次开始时。
- 连续复制失败时。
- 连续复制首次开始、失败或失败后恢复时。

复制审计事件

系统审计以下复制事件并将其记录到《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“[日志](#)”中。

- 创建、修改或销毁复制操作
- 在复制组中添加或删除共享资源
- 在目标上创建、修改、克隆、反转、断开连接或销毁复制数据包
- 创建、修改或销毁复制目标

复制和群集

可以配置从任意 ZFS Storage Appliance 到任何其他 ZFS Storage Appliance 的复制，无论每个设备是否属于群集以及该 ZFSSA 的群集对等设备是否已配置了任一方向的复制，但有以下约束：

- 不支持配置从群集的两个对等设备到同一复制目标的复制，但可以通过为同一目标 ZFSSA 使用两个不同的 IP 地址来实现类似的配置。管理员可以使用目标 ZFSSA 的多个 IP 地址在每个群集机头上创建一个复制目标，以供该机头使用。
- 当配置群集对等设备之间的复制时，请在两个控制器都处于 CLUSTERED 状态时配置复制。请勿使用专用网络地址，而是针对每个控制器的池使用单独的复制目标。

群集配置中的复制行为受限于以下规则：

- 项目和共享资源的复制更新将从已导入了包含存储池的任意群集对设备进行发送。
- 复制更新由已导入了在源的复制操作中配置的 IP 地址的任意对设备进行接收。管理员必须确保使用此 IP 地址的机头将始终拥有包含所导入副本的存储池。通过在配置群集期间将池和 IP 地址资源分配给同一个机头可确保这一点。
- 当 ZFSSA 导出相应的存储池或 IP 地址时（在接管或故障恢复过程中），正在进行的复制更新（包括到 ZFSSA 的复制更新和来自设备的复制更新）将失败。使用不受接管或故障恢复操作影响的存储池和 IP 地址的复制更新也不会受该操作影响。

有关群集和群集术语的详细信息，请查看[第 10 章 群集配置](#)。

快照和数据一致性

ZFSSA 复制快照时目标以原子方式接收每个快照，因此目标上的共享资源副本内容始终与创建快照时源上的该共享资源内容保持一致。因为在特定组中发送的所有的共享资源快照是同时创建的（请参见上文），所以成功完成复制更新后的整个数据包内容与在源上创建快照时（复制更新开始时）该组的内容完全匹配。

不过，每个共享资源的快照是单独（且依次）复制的，因此，更新数据包内某些共享资源的快照可能比更新同一数据包中其他共享资源的快照要新。在复制更新期间（在某些共享资源已更新之后但其他共享资源未更新之前）和复制更新失败后（在该复制更新之后某些共享资源可能已更新但其他共享资源可能未更新）都是如此。

总结：

- 每个共享资源在目标上始终即时一致（自身一致）。
- 当没有复制更新正在进行并且先前的复制更新已成功时，每个数据包的共享资源彼此之间也同样即时一致（数据包一致）。
- 当某个复制更新正在进行或先前的更新失败时，数据包共享资源之间可能彼此不一致，但每个共享资源仍自身一致。如果数据包一致性对应用程序十分重要，则必须克隆复制数据包，这将始终克隆最新成功接收到的每个共享资源快照。

快照管理

快照是增量复制的基础。源和目标必须始终共享一个公用快照，才能以增量方式继续复制，并且源必须知道目标具有的最新快照是哪一个。为了便于实现这一点，复制子系统将创建并管理其自己的快照。管理员通常不需要考虑这些，但由于快照对存储利用率可能有重大影响，本节还是进行了详细介绍。

特定操作的每个复制更新都包含以下步骤：

- 根据之前是否已尝试复制此操作以及目标是否已具有用于增量更新的必要快照，确定此复制更新是增量更新还是完整更新。
- 创建一个新的项目级快照。
- 发送更新。对于完整更新，将整个组的内容发送至新快照。对于增量更新，将发送先前（基础）快照与新快照之间的差异。
- 将新快照记录为下次更新的基础快照，并销毁先前的基础快照（对于增量更新）。基础快照仍然在目标上，直至获得下次更新，此时其是首先被销毁的内容。

这会对快照管理带来多种后果：

- 复制未处于主动状态时，在首次复制更新期间和初始更新后，对于在项目上或组中的任何共享资源上配置的操作，只有一个项目级快照。复制操作可以在与该操作复制（但是不在组更新过程中发送的组中）的共享资源属于同一项目的共享资源上创建快照。
- 在特定操作的后续复制更新期间，可能存在两个与该操作关联的项目级快照。如果发生了源无法确定目标是否已成功接收了新快照的故障（例如，在更新期间发生导致故障的网络中断时），在更新完成后可能会同时保留这两个快照。
- 管理员无法在不中断增量复制的情况下销毁任何与复制操作关联的快照。系统将不允许管理员销毁源或目标上增量复制所需的快照。要销毁源上的此类快照，必须销毁操作（这将销毁与操作关联的快照）。要销毁目标上的此类快照，必须先断开数据包连接（这将破坏接收对该数据包的增量更新的能力）。
- 管理员不得回滚到在任何复制快照之前创建的快照。这样做会销毁后来的复制快照，并破坏使用这些快照的所有操作的增量复制。
- 在复制时使用快照将要求使用复制的管理员了解 ZFSSA 上的 “[space management](#)” [265]，尤其是“[as it applies to snapshots](#)” [265]。

- 有关复制 LUN 的空间管理的信息，请参见[“复制 LUN 的空间管理” \[265\]](#)

复制 iSCSI 配置

如上所述，复制更新包括了在 "Shares" (共享资源) 屏幕上为项目及其共享资源指定的大多数配置。这包括与复制的 LUN 关联的任何目标组和启动器组。当使用非默认目标组和启动器组时，管理员必须确保项目内的 LUN 使用的目标组和启动器组也存在于复制目标上。组只需具有相同的名称，无需定义相同的配置。不能确保这一点可能会导致无法克隆和导出复制的 LUN。

与 LUN 关联的 SCSI GUID 将随 LUN 一起复制。因此，目标 ZFSSA 上的 LUN 将与源 ZFSSA 上的 LUN 具有相同的 SCSI GUID。不过，复制的 LUN 的克隆将拥有不同的 GUID (正如本地 LUN 的克隆与其原始 LUN 拥有不同的 GUID) 。

复制克隆

在 2009.Q3 及更早的版本中，复制仅是项目级的，且明确禁止复制下述项目：包含原始快照在项目之外的克隆的项目。由于 2010.Q1 及更高版本中具有共享资源级复制，该限制已经放宽，但管理员仍然必须考虑所复制的克隆的原始快照。特别是，克隆的初始复制会要求原始快照已经复制到目标或者在同一个更新过程中进行复制。ZFSSA 管理软件没有强制实施该限制，但是当目标上不存在原始快照时尝试复制克隆将失败。

在实际应用中，有多种方法可确保对克隆的成功复制：

- 如果克隆的原始快照在同一个项目中，只需使用项目级复制。
- 如果克隆的原始快照不在同一个项目中，或包含该原始快照的项目级复制因其他原因不合需要，请先使用共享资源级复制来复制源共享资源，然后使用项目级或共享资源级复制来复制克隆。
- 不要销毁目标系统上该克隆的原始快照，除非您也打算销毁克隆本身。

在所有情况下，原始快照的操作上的 "include snapshots" (包含快照) 属性都应当是 true，以确保原始快照已实际发送到目标。

监测复制

以下 [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Analytics”对复制可用：

- [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Data Movement Replication Operations (数据移动 – 复制操作数)”
- [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Data Movement Replication Bytes (数据移动 – 复制字节数)”

- 《Oracle ZFS Storage Appliance Analytics 指南》中的“统计信息”也可用。

复制故障

各个复制更新可能会因多种原因而失败。如有可能，ZFSSA 会在源 ZFSSA 或目标 ZFSSA 上或在失败操作的 "Replication"（复制）屏幕上发布警报，报告该故障的原因。您可以通过单击表示该操作的状态的对应橙色警报图标来获得有关故障的详细信息。以下为最常见的故障类型：

故障	详细信息
已取消	复制更新已被管理员取消。源或目标上可能已取消了复制，并且一方可能不知道另一方已取消了操作。
网络连接失败	由于网络问题，ZFSSA 无法连接到目标 ZFSSA。可能是源、目标或网络上存在错误配置。
对等设备验证失败	ZFSSA 无法验证目标的身份。当设备已重新安装或恢复了出厂设置时，此故障最常发生。要生成一组新的验证密钥，必须在源 ZFSSA 上为重新安装或恢复了出厂设置的目标配置一个新的复制目标。请参见“ 项目复制目标 ” [327]。
对等设备 RPC 失败	远程过程调用在目标系统上失败。当目标 ZFSSA 正在运行不兼容的软件时，此故障最常发生。有关详细信息，请参见“ 从 2009.Q3 及更早版本升级 ” [357]。
没有数据包	复制失败，因为目标上不存在用以包含已复制数据的数据包。由于该数据包是在配置操作时创建的，所以此故障通常发生在管理员销毁了目标上的数据包之后。如果在目标系统上未导入包含数据包的存储池，也可能发生此故障。如果池出现了故障或在目标 ZFSSA 上重新配置了网络，可能会出现未导入的情况。
存在非空数据包	复制失败，因为目标数据包包含来自先前失败的复制更新的数据。如果某个操作的首次复制更新在复制了一些数据后失败，则当尝试为该操作发送复制更新时，会发生此故障。目标 ZFSSA 不会在没有显式的管理指令的情况下销毁数据，因此它不会覆盖已部分接收的数据。管理员应当删除现有的操作和数据包，在源上创建一个新操作，然后再次启动复制。
已禁用	复制失败，因为目标上已禁用了复制。目标上已禁用了复制服务，或已经为正在复制的特定数据包禁用了复制。
目标正忙	复制失败，因为目标系统已达到最大并发复制更新数。系统限制了进行中的最大复制操作数，以避免资源耗尽。达到此限制后，后续的更新接收尝试将失败并显示此错误，后续的更新发送尝试将排入队列中，直到有资源可用。

故障	详细信息
空间不足	复制失败，因为源系统没有足够的空间来创建新快照。这可能是因为存储池中沒有可用的物理空间，或因为项目或其某个共享资源因预留空间不包括快照而将超出配额。
目标不兼容	复制失败，因为目标系统无法接收源系统的数据流格式。如果升级源系统和应用延迟更新，但不在目标上升级和应用相同更新，则可能会导致此故障。检查源系统的软件版本的发行说明以获取延迟更新列表，检查是否有某些更新对远程复制产生影响。
其他	复制失败，但源上没有提供更多信息。检查目标系统上的警报日志。如有必要，请联系技术支持以获取帮助。目前属于此类别的一些故障包括目标上没有足够的磁盘空间来接收更新，以及尝试复制目标系统上不存在其原始快照的克隆。

如果复制更新的任何部分失败，该复制更新也会失败。当前的实施会依次复制项目内的共享资源，不会回滚来自失败更新的更改。因此，当更新失败时，目标上的某些共享资源可能是最新的，而另一些共享资源则可能不是最新的。有关详细信息，请参见上文中的“快照和数据一致性”。

虽然某些数据可能已作为失败更新的一部分被成功复制，但是当前实施会重新发送已作为先前（失败的）更新的一部分发送过的所有数据。也就是说，失败的更新不会从其停止的地方开始，而是将从失败更新原来开始的位置开始。

当手动的或调度的更新失败时，系统在下次调度更新（如果有）之前不会自动重试。当连续复制失败时，系统会等待几分钟，然后重试。系统将无限地继续重试失败的连续复制。

当一个复制更新正在进行而另一更新按调度要发生时，会完全跳过后一更新，而不是在先前的更新完成后随即启动后一更新。仅当调度发生下一更新时才会发送下一更新。更新因该原因而跳过时系统会发布警报。

复制兼容性

执行复制更新之前，复制服务会验证目标系统与源系统上的新数据是否兼容。

- 如果源系统上正在使用的某些功能与目标系统不兼容且这些功能可以安全地禁用，复制服务会禁用这些功能，执行更新并发出一条警告消息。
- 如果源系统上正在使用的某些功能与目标系统不兼容且这些功能不能禁用，复制服务不会执行更新，而会发出一条错误消息。

注意：最好始终尽快升级目标。

打破复制兼容性的更新作为延迟更新提供。有关当前的列表和描述，请参见《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“更新”和 Oracle ZFS Storage Appliance 当前发行版的发行说明。

从 2009.Q3 及更早版本升级

复制实施在 2009.Q3 与 2010.Q1 版本之间已经发生重大变化。仍然强烈建议您在启动从 2009.Q3 或更低版本升级之前，暂停到 ZFSSA 的复制以及从 ZFSSA 上进行的复制。在使用滚动升级的群集中，这是强制性的。

有三个与升级到 2010.Q1 或更高版本相关的用户可见的重要更改：

- 增强了用于复制的网络协议。2009.Q3 系统可以复制到运行任何版本（包括 2010.Q1 和更高版本）的系统，而运行 2010.Q1 或更高版本的系统只能复制到运行 2010.Q1 或更高版本的其他系统。在实际应用中，这意味着复制目标必须与其复制源同时升级或在其之前升级，以避免因协议版本不兼容而导致失败。
- 复制操作配置现在存储在存储池中，而不是在机头系统上。因此，在从 2009.Q3 或更早版本升级到 2010.Q1 之后，管理员必须应用延迟更新来迁移其复制配置。
- * 在应用这些更新之前，所传入的对现有副本的复制更新将失败，且不会为在 2009.Q3 或更早版本中配置的操作发送复制更新。此外，对于无法从 BUI 或 CLI 中管理的未迁移副本，将使用存储池中的空间。
- * 在应用这些更新之后，与所有的延迟更新一样，对系统软件进行回滚将产生不确定的结果。应当预料到在较旧的版本中，复制的数据将不可访问，将取消所有复制操作的配置，传入的复制更新将是完整更新。
- 复制授权已从其自己的范围移动到了 "Projects and Shares"（项目和共享资源）范围中。在 2009.Q3 或更早版本中配置的任何复制授权在 2010.Q1 中将不再存在。对复制使用细粒度访问控制的管理员应在升级后将新的复制授权委托给相应的管理员。

影子迁移

本章介绍 ZFSSA 的影子迁移。

数据迁移

管理员的一个常见任务是将数据从一个位置移至另一个位置。从最抽象的意义上讲，此问题包含大量使用案例，从在服务器之间复制数据到将手提电脑上的用户数据与服务器保持同步。可以使用许多外部工具执行此操作，但是 ZFSSA 具有两个集成解决方案，用于迁移处理最常用案例的数据。第一个是[第 13 章 复制](#)，用于在一个或多个 ZFSSA 之间复制数据，这将单独讲述。第二个是影子迁移，在此处进行讲述。

影子迁移是从外部 NAS 源迁移数据的过程，并且会在迁移完成后替换或停止使用原始数据。这种迁移最常用在将某一新 ZFSSA 加入现有环境以接替另一服务器承担的文件共享职责的情况下，但也可能有其他一些新颖的用途，如下文所述。

传统数据迁移

传统文件迁移通常采用以下两种方式之一进行：反复同步或外部插入。

通过同步进行迁移

这种方法使用活动的主机 X 向新主机 Y 迁移数据，同时主机 X 仍然保持活动状态。在迁移过程中，客户机仍然对原始主机进行读写操作。在进行初次迁移后，将反复地发送增量更改数据，一直到增量数据少到足以在一个停机时段内发送完毕为止。此时，原始共享资源将变为只读状态，最后一部分增量数据将发送到新主机，而所有客户机将更新为指向该新位置。实现这种迁移的最常用方法是通过 rsync 工具，但是还有其他集成的工具可用。这种方法有几个缺点：

- 尽管预期的停机时间较短，但不易量化。如果用户在计划的停机时间即将到来之前提交了大量更改数据，将会延长停机时段。
- 在迁移期间，新服务器处于空闲状态。因为新服务器往往具备新的功能或性能上的改进，所以，这会在可能很长的迁移期间内造成资源浪费。

- 在多个文件系统之间协调的任务很繁重。在迁移数十个或数百个文件系统时，每次迁移都需要不同长短的时间，因此，必须在所有文件系统构成的联合体内计划好各自的停机时间。

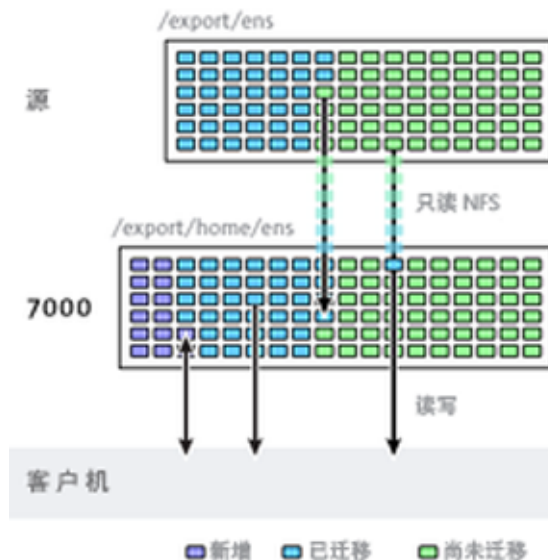
通过外部插入进行迁移

这种方法使用活动的主机 X，并另外加入一个新 ZFSSA M，由后者将数据迁移到新主机 Y。所有客户机会立即更新为指向 M，而数据自动在后台进行迁移。这可以为迁移选项带来更多灵活性（例如，能够将来在不停机情况下向新服务器进行迁移），并可以利用新服务器来处理已迁移的数据，但也有一些显著的缺点：

- 迁移 ZFSSA 作为新的物理机，产生相关成本（初始投资、支持成本、供电和冷却等）及其他管理开销。
- 迁移 ZFSSA 成为系统中新的故障点。
- 迁移 ZFSSA 插入已迁移的数据，这样会产生额外的延迟，通常是永久性的。这些迁移 ZFSSA 通常需要持续运行，当然也可以为其另外计划停机时段并将其停用。

影子迁移

图 14-1 影子迁移



影子迁移需要使用插入，但是已集成到 ZFSSA 本身之内，因此不需要另外的物理机。当创建共享资源后，它们可以选择在本地或在 NFS 上充当现有目录的“影子”。在这种方案中，将一次性计划好停机时间，让源 ZFSSA X 进入只读模式，并创建设置了影子属性的共享资源，而客户机将更新为指向 Sun Storage 7000 ZFSSA 上的新共享资源。然后，客户机可以以读写模式访问该 ZFSSA。

在设置了影子属性后，数据将在后台从源 ZFSSA 本地透明地进行迁移。如果客户机请求访问某个尚未迁移的文件，源 ZFSSA 会自动将该文件迁移到本地服务器，然后响应请求。对于某些客户机请求，这会产生一些初始延迟，但一旦将文件迁移完毕，所有访问都可在 ZFSSA 本地完成，从而能完全发挥 ZFSSA 本身的性能。当文件系统的当前工作集远远小于总数据大小时，就经常出现这种情况，因此一旦将工作集迁移完毕，无论源上的总数据大小如何，都不会感觉到对性能造成影响。

影子迁移的缺点是需要在完成数据迁移之前执行一些相关任务，不过只有插入方法是这样。在迁移过程中，数据的某些部分存在于两个位置，这意味着备份更加复杂，快照可能不完整且（或）仅存在于一个主机上。因此，非常重要的一点是，两个主机之间进行的任何迁移都要首先进行彻底测试，确保身份管理和访问控制设置正确。这并不需要测试整个数据迁移过程，但应该验证非全局可读的文件或目录是否能正确迁移、ACL（如果存在）是否能保留，以及身份是否能在新系统中正确表示。

影子迁移是使用文件系统中的磁盘数据实施的，因此没有外部数据库，也不在存储池外部的本地存储数据。如果存储池在群集中进行了故障转移，或者两个系统磁盘都出现故障而需要有新的机头节点时，不中断地进行影子迁移所需的所有数据将随该存储池一起保存。

影子迁移的行为方式

影子源的限制

- 要正确地迁移数据，源文件系统或目录必须为只读状态。对文件源所做的更改是否会传播取决于其发生的时机，对目录结构的更改可能会导致 ZFSSA 上发生无法恢复的错误。
- 影子迁移只支持从 NFS 源进行迁移。NFSv4 共享资源将产生最佳效果。也可以进行 NFSv2 和 NFSv3 迁移，但会在迁移过程中丢失 ACL，且相对于 NFSv2 过大的文件无法使用该协议进行迁移。不支持从 SMB 源进行迁移。
- 不支持 LUN 的影子迁移。

迁移过程中的影子文件系统语义

如果客户机访问尚未进行迁移的文件或目录，将对文件系统产生以下可监测的影响：

- 对于目录访问，客户机请求将被阻止，直到整个目录完成迁移为止。对于文件访问，将只迁移所请求的文件部分，多个客户机可以同时迁移文件的不同部分。
- 可以在影子文件系统中任意地对文件和目录进行重命名、删除或覆盖，而对迁移过程没有任何影响。
- 对于作为硬链接的文件，在迁移完成之前，硬链接计数可能与源文件不匹配。
- 大多数文件属性都会在创建目录时进行迁移，但只有在对文件完成读取或写入操作后，盘上大小（UNIX stat 结构中的 st_nblocks）才可用。逻辑大小是正确的，但 du(1) 或其他命令会报告大小为零，直到文件内容已实际迁移为止。
- 重新引导 ZFSSA 后，迁移将从最初停止的位置开始继续进行。虽然不需要重新迁移数据，但可能需要遍历本地文件系统中某些已迁移的部分，因此，可能会由于中断而对总迁移时间有所影响。
- 数据迁移会利用文件上的专用扩展属性。通常，这些属性是无法查看的，除非它们位于文件系统的根目录上，或者通过快照来查看。如果添加、修改或删除以 SUNWshadow 开头的任何扩展属性，则会对迁移过程产生不确定的影响，并导致不完整或已损坏状态。此外，文件系统级别的状态存储在文件系统根目录的 .SUNWshadow 目录中。如果对此内容进行任何修改，则会产生类似的影响。
- 文件系统完成迁移后，将发布一个警报，同时会删除影子属性以及任何适用的元数据。之后，该文件系统就与标准文件系统没有任何差别了。
- 可以通过使用 NFSv4 自动客户机挂载（有时称为“镜像挂载”）或嵌套本地挂载将多个文件系统的数据迁移到单个文件系统中。

标识和 ACL 迁移

要正确迁移文件的标识信息，必须遵循以下规则：

- 迁移源和目标 ZFSSA 必须具有相同的名称服务配置。
- 迁移源和目标 ZFSSA 必须具有相同的 NFSv4 mapid 域
- 迁移源必须支持 NFSv4。可以使用 NFSv3，但会丢失某些信息。基本标识信息（所有者和组）和 POSIX 权限将保留下来，但所有 ACL 都会丢失。
- 必须以 ZFSSA 的 root 权限导出迁移源。

如果您发现文件或目录的所有者为 "nobody"，则可能的原因是，该 ZFSSA 未正确设置名称服务，或者 NFSv4 mapid 域不同。如果在遍历客户机本应能够访问的文件系统时显示 "permission denied"（权限被拒绝）错误，则最可能的原因是，没有以 root 权限导出迁移源。

影子迁移管理

创建影子文件系统

只有在创建文件系统后，才能设置影子迁移源。在 BUI 中，可以通过文件系统创建对话框来创建文件系统。在 CLI 中，可以通过 `shadow` 属性来创建文件系统。此属性采用下列形式之一：

- 本地 – `file:///<path>`
- NFS – `nfs://<host>/<path>`

在 BUI 中，对于 NFS 挂载，还可以使用备用形式 `<host>:/<path>`，此形式与 UNIX 系统中使用的语法是一致的。通过 BUI，还可以使用下拉菜单来设置此设置的协议部分（`file://` 或 `nfs://`）。创建文件系统时，服务器将验证该路径是否存在，并且是否可挂载。

管理后台迁移

创建共享资源后，该共享资源将自动开始在后台迁移，并为内嵌请求提供支持。此迁移通过“影子迁移服务” [216] 进行控制。此服务提供了一个全局可调参数，即，此任务专用的线程数。提高此线程数，会提高并行性，但会占用更多资源。

可以禁用影子迁移服务，但只能出于测试目的来禁用，或者只能在正在进行的影子迁移使系统负载过重，以至于需要临时停止时禁用。禁用影子迁移服务后，仍会根据需要迁移同步请求，但不会进行后台迁移。禁用此服务后，任何影子迁移都将无法完成，即使已手动读取文件系统的所有内容。强烈建议始终启用此服务。

处理迁移错误

由于影子迁移要求将新写入的数据提交到服务器后，才能完成迁移，因此，必须对迁移进行测试并监视是否出现任何错误，这一点很重要。后台迁移期间遇到的错误会在 BUI 中显示为影子迁移状态的一部分。不会跟踪在其他同步迁移期间遇到的错误，这些错误将在后台进程访问受影响的文件时进行说明。对于每个文件，系统会保留远程文件名以及特定的错误。单击错误计数旁边的信息图标将显示该详细列表。修复错误后，该错误列表不会更新，而是会在成功完成迁移后被清除。

只有成功迁移所有文件后，影子迁移才会完成。如果出现错误，则后台迁移将不断重试，直到成功为止。这样可以使管理员修复任何错误（例如，权限问题），然后完成迁移，并确保迁移成功。如果由于持续出现错误而无法完成迁移，则可以取消该迁移，从而在本地文件系统中保留可迁移的数据。而取消迁移只应作为最后一项措施来使用，一旦取消迁移，该迁移将无法再继续。

监视迁移进度

在给定的迁移运行环境下，监视影子迁移的进度比较困难。一个文件系统可以作为一个文件系统全部内容或部分内容的影子，也可以作为具有嵌套挂载点的多个文件系统的影子。因此，无法请求有关源的统计信息，并确保这些信息是正确的。此外，即使迁移单个文件系统，用于计算可用大小的方法在各个系统之间也是不一致的。例如，远程文件系统可能会使用压缩，或者可能包括元数据开销，也可能不包括元数据开销。因此，对于任何特定的迁移，无法显示准确的进度条。

此 ZFSSA 会提供以下信息，这些信息可确保准确无误：

- 到目前为止，本地文件系统的本地大小
- 到目前为止，所复制数据的逻辑大小
- 到目前为止，迁移数据所花费的时间

这些值是通过标准文件系统属性以及影子迁移节点（或 UI 面板）的属性在 BUI 和 CLI 中提供的。如果您知道远程文件系统的大小，则可以使用此大小来估计进度。复制的数据的大小仅包含需要从源迁移的纯文件内容。目录、元数据和扩展属性不包括在此计算中。虽然到目前为止迁移的数据的大小仅包括远程迁移的数据，但是，恢复后台迁移可能会遍历已完成迁移的文件系统部分。这可能会导致在处理这些初始目录时运行相当快速，但当到达尚未迁移的文件系统部分后，运行速度会减缓。

虽然对进度没有准确的测量方法，但是，ZFSSA 会根据相对统一目录树的假设，尝试对剩余的数据量进行估计。此估计值的范围可能相当准确，也可能毫无意义，具体取决于数据集，因此该信息仅供参考。例如，一个迁移可能具有相对较浅的文件系统树，但在最后访问的单个目录中却存在大量数据。在这种情况下，迁移将显示为接近完成，然后在搜索此新树时，完成进度会迅速下降到一个非常小的百分比。相反，如果首先处理的是大型目录，则在进行估计时可能会假定所有其他目录都同样具有大量的数据，而在发现这些目录大部分为空时，估计值会快速从一个小的百分比上升到接近完成。测量进度的最佳方法是设置一个测试迁移，让它运行至完成，然后使用该值来估计具有类似布局和大小文件系统的进度。

取消迁移

迁移可以取消，但只有在出现源不再可用这种极端情况时才应取消迁移。一旦取消迁移，将无法恢复。主要目的是当源中存在无法更正的错误时允许迁移完成。如果除了少数文件或目录之外，整个文件系统已完成迁移，并且无法更正这些错误（例如，源已永久损坏），则取消迁移将允许本地文件系统将状态恢复为“正常”文件系统。

要在 BUI 中取消迁移，请在存在问题的共享资源的左列中，单击进度条旁边的关闭图标。在 CLI 中，先迁移到文件系统下的 shadow 节点，然后运行 cancel 命令。

影子文件系统的快照

可以创建影子文件系统的快照，不过，快照中所包含的对象的状态是任意的。尚未迁移的文件不会出现，并且在快照中可能会显示实施详细信息（例如 SUNWshadow 扩展属性）。此快照可用来恢复自原始迁移开始后已迁移或修改的各个文件。因此，建议将所有快照保存在源中直到迁移完成，以便在需要时仍可从源检索未迁移的文件。根据保留策略，可能需要在源中扩展保留才能满足服务要求。

虽然可以创建快照，但这些快照无法回滚到也不能成为克隆的源。这反映了在迁移期间磁盘上数据的不一致状态。

备份影子文件系统

与任何其他系统一样，可以使用 NDMP 来备份正在迁移影子数据的文件系统。影子设置将随备份流保留，但是只有当执行文件系统的完整恢复且共享资源事先不存在时才能恢复。从这样的备份流恢复各个文件或恢复到现有文件系统可能会导致状态不一致或数据损坏。在完整文件系统恢复期间，文件系统也将处于不一致的状态（除部分恢复的常见不一致之外的不一致），并且影子迁移将处于非活动状态。只有在恢复完成时才能恢复影子设置。如果影子源不再存在或已移动，则管理员可以查看任何错误，并根据需要对其进行更正。

复制影子文件系统

使用常规机制可以复制正在迁移影子数据的文件系统，但在数据流中只会发送迁移的数据。因此，远程端仅包含部分数据，这些数据可能表现出不一致的状态。影子设置随复制流一同发送，因此当对远程目标进行故障转移时，将保留相同的影子设置。与恢复 NDMP 备份流一样，此设置在远程目标的上下文中可能不正确。对目标进行故障转移后，管理员可以查看任何错误，并根据需要为新环境更正影子设置。

影子迁移分析

除了以每个共享资源为标准监视外，还可以通过 [《Oracle ZFS Storage Appliance Analytics 指南》](#) 中的“Analytics”在系统范围内监视影子迁移。影子迁移分析信息是在“Data Movement”（数据移动）类别下提供的。提供了以下两种基本统计信息：

影子迁移请求数

此统计信息跟踪针对未缓存且已知在文件系统本地的文件或目录的影子迁移请求数。此统计信息同时考虑了已迁移的和未迁移的文件和目录，且可用于跟踪在影子迁移过程中

发生的延迟，以及跟踪后台迁移的进度。它可以按文件、共享资源、项目或延迟进行细分。由于它当前同时包含同步和异步（后台）迁移，因此无法仅查看对客户机可见的延迟。

影子迁移字节数

此统计信息跟踪在迁移文件或目录内容期间传送的字节数。此统计信息不适用于元数据（扩展属性、ACL 等等）。此统计信息给出了所传送的数据的粗略近似值，但包含大量元数据的源数据集将显示不成比例的小带宽。可以通过查看网络 Analytics 信息来监测整个带宽。此统计信息可以按本地文件名、共享资源或项目进行细分。

影子迁移操作数

此统计信息跟踪需要转到源文件系统的操作数。此统计信息可用于跟踪来自影子迁移源的请求的延迟。它可以按文件、共享资源、项目或延迟进行细分。

迁移本地文件系统

除了其主要目的是从远程源迁移数据外，同一机制还可用于将数据从本地文件系统迁移到 ZFSSA 中的其他文件系统。此机制可用于更改无法通过其他方法修改的设置，例如创建文件系统的压缩版本，或在该操作后更改文件系统的记录大小。在此模型中，会将旧的共享资源（或共享资源内的子目录）设置为只读的或将其移到一边，然后使用 file 协议在设置了影子属性的情况下创建一个新的共享资源。客户机访问此新共享资源，并且使用新共享资源的设置写入数据。

影子迁移任务

在尝试完整的迁移之前，请务必对迁移进行测试以确保 ZFSSA 具有合适的权限且已正确转换了安全性属性。在确认基本设置正常发挥作用后，可以对共享资源进行设置以便执行最终迁移。

▼ 测试潜在的影子迁移

1. 对源进行配置以便 ZFSSA 对共享资源具有 root 访问权限。这通常涉及添加基于 NFS 主机的例外，或者设置匿名用户映射（后者具有更重大的安全隐患）。
2. 在本地文件系统中创建一个共享资源，并在 CLI 中将 shadow 属性设置为 'nfs://<host>/<snapshotpath>' 或者在 BUI 中将 shadow 属性设置为 '<host>/<snapshotpath>'（选择

'NFS' 协议)。快照应当是源的只读副本。如果没有快照可用，则可以使用读写源，但可能会导致不确定的错误。

3. 通过遍历文件结构验证正确保留了文件内容和身份映射。
4. 如果数据源是只读的（与快照一样），请让迁移完成并确认在传送中没有错误。

▼ 从活动 NFS 服务器迁移数据

1. 安排停机时间，可以在该时间段内暂停客户机并将其重新配置为指向新服务器。
2. 对源进行配置以便 ZFSSA 对共享资源具有 root 访问权限。这通常涉及添加基于 NFS 主机的例外，或者设置匿名用户映射（后者具有更重大的安全隐患）。
3. 将源配置为只读的。此步骤从技术角度来讲是可选的，但如果错误配置的客户机在迁移过程中无法向源进行写入，则执行此步骤更容易保证符合性。
4. 在本地文件系统上创建一个共享资源，并在 CLI 中将 shadow 属性设置为 'nfs://<host>/<path>' 或者在 BUI 中将 shadow 属性设置为 '<host>/<path>'（选择 'NFS' 协议）。
5. 重新配置客户机以指向 SS7000 上的本地共享资源。

此时，影子迁移应在后台运行，且系统可以根据需要为客户机请求提供服务。您可以如上所述来监测进度。通过 CLI 脚本化，可以在计划的单个停机时间内创建多个共享资源。

◆◆◆ 第 15 章

CLI 脚本化

设计 CLI 旨在为执行重复性任务提供一个强大的脚本化环境。

自动化访问

您可使用“[批处理命令](#)” [369]或“[脚本化命令](#)” [370]（或某种组合），但在任何情况下自动化基础结构都需要对设备进行自动化访问。必须通过[第 7 章 用户配置](#)、“[用户授权](#)” [128]和“[使用 CLI 设置 SSH 公钥](#)” [138]进行。

批处理命令

最简单的脚本化机制是批处理设备 Shell 命令。例如，要在项目 "myproj" 和文件系统 "myfs" 中自动获取名为 "newsnap" 的快照，请在文件中使用下面的命令：

```
shares
select myproj
select myfs
snapshots snapshot newsnap
```

然后使用 ssh 命令传输到设备，将标准输入重定向到该文件：

```
% ssh root@dory < myfile.txt
```

在许多 Shell 中，可以使用“Here 文件”简化此过程，在其中将输入（甚至是标记）发送到标准输入。如果改用 Here 文件，上例应使用以下内容：

```
% '''ssh root@dory << EOF
shares
select myproj
select myfs
snapshots snapshot newsnap
EOF'''
```

此机制对于最简单的自动化类型已经足够，如果在客户机上使用较高级别的 Shell 脚本语言封装在编程逻辑中，可能也足够，但是总的来说，还有很多可以改善的地方。

脚本化命令

批处理命令对于最简单的操作已经足够，但是要封装在编程逻辑中，可能有点麻烦。例如，如果要获取有关每个共享资源空间使用情况的信息，则必须执行许多不同的 CLI 的调用，并在解析特定命令输出的客户机上使用较高级别的语言封装。这会导致自动化基础结构缓慢而不稳定。为了实现更快且最稳定的自动化，设备提供了基于 ECMAScript 3 的丰富脚本化环境。有关 ECMAScript 的说明不在本文档的范围内，但是它是一种动态类型化的语言，使用类似于 C 的语法，以允许：

- 条件代码流 (if/else)
- 迭代代码流 (while、for 等)
- 结构化和数组数据处理（通过先进的对象和数组类型实现）
- 类似于 Perl 的正则表达式和字符串处理 (split()、join()等)
- Exceptions（例外）
- 复杂的函数式语言功能，例如闭包

脚本环境

在 CLI 中，使用 script 命令进入脚本环境：

```
dory:> script
("." to run)>
```

出现脚本环境提示符时，即可输入脚本，最后另起一行单独输入 "." 来执行该脚本：

```
dory:> script
("." to run)> for (i = 10; i > 0; i--)
("." to run)>   printf("%d... ", i);
("." to run)> printf("Blastoff!\n");
("." to run)> .
10... 9... 8... 7... 6... 5... 4... 3... 2... 1... Blastoff!
```

如果脚本只有一行，则可以直接将其作为参数提供给 script 命令，以便轻松地浏览脚本：

```
dory:> script print("It is now " + new Date())
It is now Tue Oct 14 2009 05:33:01 GMT+0000 (UTC)
```

与系统交互

当然，除非可以与整个系统交互，否则脚本的用处不大。下面是一些允许脚本与系统交互的内置函数：

表 15-1 支持系统交互的内置函数

函数	说明
get	获取指定的属性的值。请注意，此函数以原生格式返回值，例如将日期作为 Date 对象返回。
list	返回对应于当前上下文动态子代的标记数组。
run	运行 Shell 中指定的命令，并将所有输出作为字符串返回。请注意，如果输出包含多行，则返回的字符串将包含嵌入的换行符。
props	返回当前节点的属性名称数组。
set	采用两个字符串参数，将指定的属性设置为指定的值。
choicies	返回一组值已知且可枚举的任何属性的有效属性值数组。

Run 函数

脚本与较大系统交互的最简单方式就是使用 "run" 函数：它获取要运行的命令，并将该命令的输出作为字符串返回。例如：

```
dory:> configuration version script dump(run('get boot_time'))
'                               boot_time = 2009-10-12 07:02:17\n'
```

内置的 dump 函数转储参数，但不展开任何嵌入的换行符。ECMAScript 的字符串处理工具可用于拆分输出。例如，按空格拆分上面的结果：

```
dory:> configuration version script dump(run('get boot_time').split(/\s+/))
[&#39;;', 'boot_time', '=', '2009-10-12', '07:02:17', &#39;']
```

Get 函数

run 函数非常强大，它往往专门依靠解析输出来获取有关系统的信息 – 但是这显然存在缺陷，因为它让脚本解析用户可读的输出，输出在将来可能会发生更改，也可能不会。要更稳妥地收集有关系统的信息，请使用内置的 "get" 函数。对于 boot_time 属性，此函数返回的内容不是字符串，而是 ECMAScript Date 对象，以允许通过编程方式处理属性值。例如，您可能希望结合使用 boot_time 属性和当前时间确定自引导以来经过的时间：

```
script
run('configuration version');
now = new Date();
uptime = (now.valueOf() - get('boot_time').valueOf()) / 1000;
printf('up %d day%s, %d hour%s, %d minute%s, %d second%s\n',
    d = uptime / 86400, d < 1 || d >= 2 ? 's' : '',
    h = (uptime / 3600) % 24, h < 1 || h >= 2 ? 's': '',
    m = (uptime / 60) % 60, m < 1 || m >= 2 ? 's': '',
```

```
s = uptime % 60, s < 1 || s >= 2 ? 's': ''');
```

假设上面的内容保存为 "uptime.aksh", 则可以通过以下方式运行：

```
% ssh root@dory < uptime.aksh
Pseudo-terminal will not be allocated because stdin is not a terminal.
Password:
up 2 days, 10 hours, 47 minutes, 48 seconds
```

由于是 ssh 客户机，因此会出现有关伪终端分配的消息；该消息所指的问题可以通过为 ssh 指定 "-T" 选项进行处理。

List 函数

在具有动态子代的上下文中，以编程方式迭代这些子代可能非常有用。这可以使用 list 函数实现，该函数返回动态子代数组。例如，下面是迭代每个项目中每个共享资源的脚本，输出所用空间和可用空间：

```
script
  run('shares');
  projects = list();

  for (i = 0; i < projects.length; i++) {
    run('select ' + projects[i]);
    shares = list();

    for (j = 0; j < shares.length; j++) {
      run('select ' + shares[j]);
      printf("%s/%s %1.64g %1.64g\n", projects[i], shares[j],
        get('space_data'), get('space_available'));
      run('cd ..');
    }

    run('cd ..');
  }
}
```

下面是运行该脚本的输出（假设将其保存到名为 "space.aksh" 的文件中）：

```
% ssh root@koi < space.aksh
Password:
admin/accounts 18432 266617007104
admin/exports 18432 266617007104
admin/primary 18432 266617007104
admin/traffic 18432 266617007104
admin/workflow 18432 266617007104
aleventhal/hw_eng 18432 266617007104
bcantrill/analytx 1073964032 266617007104
bgregg/dashbd 18432 266617007104
bgregg/filesys01 26112 107374156288
bpijewski/access_ctrl 18432 266617007104
...
```

如果有人希望获得以上结果的 "易于阅读"（虽然通过编程方式比较难处理）版本，则可以直接解析 get 命令的输出：

```

script
    run('shares');
    projects = list();

    printf('%-40s %-10s %-10s\n', 'SHARE', 'USED', 'AVAILABLE');

    for (i = 0; i < projects.length; i++) {
        run('select ' + projects[i]);
        shares = list();

        for (j = 0; j < shares.length; j++) {
            run('select ' + shares[j]);

            share = projects[i] + '/' + shares[j];
            used = run('get space_data').split(/\s+/)[3];
            avail = run('get space_available').split(/\s+/)[3];

            printf('%-40s %-10s %-10s\n', share, used, avail);
            run('cd ..');
        }

        run('cd ..');
    }
}

```

以下是运行该新脚本的部分输出（假设将其命名为 "prettyspace.aksh"）：

```

% ssh root@koi < prettyspace.aksh
Password:
SHARE                                USED      AVAILABLE
admin/accounts                      18K       248G
admin/exports                       18K       248G
admin/primary                       18K       248G
admin/traffic                       18K       248G
admin/workflow                      18K       248G
aleventhal/hw_eng                  18K       248G
bcantrill/analytx                  1.00G     248G
bgregg/dashbd                     18K       248G
bgregg/filesys01                   25.5K     100G
bpijewski/access_ctrl              18K       248G
...

```

Children 函数

即使是在具有静态子代的上下文中，以编程方式迭代这些子代同样非常有用。这可以使用 `children` 函数实现，该函数返回静态子代数组。例如，下面是迭代每个服务的脚本，输出服务的状态：

```

configuration services
script
    var svcs = children();
    for (var i = 0; i < svcs.length; ++i) {
        run(svcs[i]);
        try {
            printf("%-10s %s\n", svcs[i], get('<status>'));
        } catch (err) { }
    }
}

```

```

        run("done");
    }

```

下面是运行该脚本的输出（假设将其保存到名为 "svcinfo.aksh" 的文件中）：

```

% ssh root@koi < space.aksh
Password:
cifs      disabled
dns       online
ftp       disabled
http      disabled
identity  online
idmap     online
ipmp      online
iscsi     online
ldap      disabled
ndmp      online
nfs       online
nis       online
ntp       online
scrk      online
sftp      disabled
smtp      online
snmp      disabled
ssh       online
tags      online
vscan     disabled

```

Choices 函数

Choices 函数将返回一组值已知且可枚举的任何属性的有效属性值数组。例如，以下脚本将使用 choices 函数检索共享资源节点上所有池的列表，然后迭代所有池来列出项目和共享资源以及可用空间。

```

fmt = '%-40s %-15s %-15s\n';
printf(fmt, 'SHARE', 'USED', 'AVAILABLE');
run('cd /');
run('shares');
pools = choices('pool');
for (p = 0; p < pools.length; p++) {
    set('pool', pools[p]);
    projects = list();
    for (i = 0; i < projects.length; i++) {
        run('select ' + projects[i]);
        shares = list();
        for (j = 0; j < shares.length; j++) {
            run('select ' + shares[j]);
            share = pools[p] + ':' + projects[i] + '/' + shares[j];
            printf(fmt, share, get('space_data'),
                get('space_available'));
            run('cd ..');
        }
        run('cd ..');
    }
}
}

```

以下是运行该脚本的输出：

SHARE	USED	AVAILABLE
pond:projectA/fs1	31744	566196178944
pond:projectA/fs2	31744	566196178944
pond:projectB/lun1	21474836480	587670999040
puddle:deptA/share1	238475	467539219283
puddle:deptB/share1	129564	467539219283
puddle:deptB/share2	19283747	467539219283

生成输出

报告系统的状态需要生成输出。脚本具有多个可供其生成输出的内置函数：

表 15-2 生成输出的内置函数

函数	说明
dump	将指定的参数转储到终端，但不展开嵌入的换行符。对象将以类似 JSON 的格式显示。对于调试很有用。
print	将指定的对象作为字符串输出，后跟一个换行符。如果对象没有 toString 方法，则以不透明的方式输出。
printf	与 C 语言的 printf(3C) 一样，根据指定的格式字符串输出指定的参数。

处理错误

当生成错误时，会引发异常。异常通常是包含以下成员的对象：

- code – 与错误关联的数字代码
- message – 与错误关联的用户可读消息

可以捕获并处理异常，否则可能会在脚本环境中引发。如果脚本环境拥有未捕获的异常，则 CLI 将显示详细信息。例如：

```
dory:> script run('not a cmd')
error: uncaught error exception (code EAKSH_BADCMD) in script: invalid command
      "not a cmd" (encountered while attempting to run command "not a cmd")
```

可以通过捕获并转储异常来了解有关该异常的更多详细信息：

```
dory:> script try { run('not a cmd') } catch (err) { dump(err); }
{
  toString: <function>,
  code: 10004,
```

```
        message: 'invalid command "not a cmd" (encountered while attempting to
                    run command "not a cmd")'
    }
}
```

这还使您拥有丰富的错误处理方法，例如：

```
#!/usr/bin/ksh -p

ssh -T root@dory <<EOF
script
    try {
        run('shares select default select $1');
    } catch (err) {
        if (err.code == EAKSH_ENTITY_BADSELECT) {
            printf('error: "$1" is not a share in the ' +
                'default project\n');
            exit(1);
        }

        throw (err);
    }

    printf('default/$1: compression is %s\n', get('compression'));
    exit(0);
EOF
```

如果该脚本的名称为 "share.ksh" 并使用无效的共享资源名称运行，则将生成丰富的错误消息：

```
% ksh ./share.ksh bogus
error: "bogus" is not a share in the default project
```


◆◆◆ 第 16 章

维护工作流

工作流是上载到 ZFSSA 并由 ZFSSA 本身管理的[第 15 章 CLI 脚本化](#)。工作流可以在浏览器界面或命令行界面中通过先进的方式参数化和执行。可以选择将工作流作为[第 9 章 警报配置](#)执行，或者在指定的时间执行。因此，工作流允许 ZFSSA 以捕获特定策略和过程的方式进行扩展，并可用于（例如）正式对特定组织或应用程序的最佳做法进行编码。

使用工作流

工作流包含在有效的 ECMAScript 文件中，包含一个全局变量 `workflow`。这是一个必须至少包含三个成员的对象：

表 16-1 必要的对象成员

必要成员	类型	说明
<code>name</code>	字符串	工作流的名称
<code>description</code>	字符串	工作流的说明
<code>execute</code>	函数	执行工作流的函数

下面是规范的简单工作流：

```
var workflow = {
  name: 'Hello world',
  description: 'Bids a greeting to the world',
  execute: function () { return ('hello world!') }
};
```

上载此工作流将生成一个名为 "Hello world" 的新工作流；执行该工作流将输出 "hello world!"

工作流执行上下文

工作流在 ZFSSA Shell 中异步执行，（默认）在用户执行工作流时运行。因此，工作流可自行使用[第 15 章 CLI 脚本化](#)，并且可能会与 ZFSSA 进行交互，就像 ZFSSA Shell 的任何其他实例一样。即工作流可以执行命令、解析输出、修改状态等等。下面是一个使用 run 函数返回当前 CPU 利用率的比较复杂的示例：

```
var workflow = {
  name: 'CPU utilization',
  description: 'Displays the current CPU utilization',
  execute: function () {
    run('analytics datasets select name=cpu.utilization');
    cpu = run('csv 1').split('\n')[1].split(',');
    return ('At ' + cpu[0] + ', utilization is ' + cpu[1] + '%');
  }
};
```

工作流参数

不处理输入的工作流的作用域是有限的；许多工作流需要参数化才能使用。这可以通过将 parameters 成员添加到全局 workflow 对象来实现。而 parameters 成员又是应将每个参数作为一个成员的对象。每个 parameters 成员必须具有以下成员：

表 16-2 必要的工作流参数成员

必要成员	类型	说明
label	字符串	指示工作流参数输入的标签
type	字符串	工作流参数的类型

type 成员必须设置为以下类型之一：

表 16-3 成员类型名称

类型名称	说明
Boolean	布尔值
ChooseOne	一系列指定值中的一个
EmailAddress	电子邮件地址
File	要传输到 ZFSSA 的文件
Host	有效主机，名称或点分十进制

类型名称	说明
HostName	有效主机名
HostPort	有效的可用端口
Integer	整数
NetAddress	网络地址
NodeName	网络节点的名称
NonNegativeInteger	大于或等于零的整数
Number	任意数字 – 包括浮点
Password	密码
Permissions	POSIX 权限
Port	端口号
Size	大小
String	字符串
StringList	字符串列表

根据指定的类型，执行工作流时会生成相应的输入格式。例如，下面的工作流具有两个参数，业务部门的名称（要用作项目）和共享资源的名称（要用作共享资源名称）：

```
var workflow = {
  name: 'New share',
  description: 'Creates a new share in a business unit',
  parameters: {
    name: {
      label: 'Name of new share',
      type: 'String'
    },
    unit: {
      label: 'Business unit',
      type: 'String'
    }
  },
  execute: function (params) {
    run('shares select ' + params.unit);
    run('filesystem ' + params.name);
    run('commit');
    return ('Created new share "' + params.name + '"');
  }
};
```

如果上载此工作流并执行，则系统将显示一个对话框，提示输入共享资源和业务部门的名称。创建共享资源后，将生成一条消息，指示相应信息。

约束参数

对于某些参数，有人不希望允许使用任意字符串，而是希望将输入限制为少量备选选项之一。这些参数应该指定为 ChooseOne 类型，包含参数的对象必须具有另外两个成员：

表 16-4 约束参数的必要成员

必要成员	类型	说明
options	数组	一个字符串数组，指定有效的选项
optionlabels	数组	一个字符串数组，指定与在 options 中所指定的选项相关联的标签

使用 ChooseOne 参数类型，我们可以强化前面的示例，将业务部门限制为少量预定义的值之一：

```
var workflow = {
  name: 'Create share',
  description: 'Creates a new share in a business unit',
  parameters: {
    name: {
      label: 'Name of new share',
      type: 'String'
    },
    unit: {
      label: 'Business unit',
      type: 'ChooseOne',
      options: [ 'development', 'finance', 'qa', 'sales' ],
      optionlabels: [ 'Development', 'Finance',
        'Quality Assurance', 'Sales/Administrative' ],
    }
  },
  execute: function (params) {
    run('shares select ' + params.unit);
    run('filesystem ' + params.name);
    run('commit');
    return ('Created new share "' + params.name + '"');
  }
};
```

当执行此工作流时，unit 参数不能手动输入 – 它将从指定的可用选项列表中进行选择。

可选参数

有些参数可以视为可选参数，因为 UI 不强制要求将这些参数设置为任何值即允许工作流执行。此类参数通过 parameters 成员的 optional 字段指明：

表 16-5
 可选参数的必要成员

可选成员	类型	说明
optional	Boolean（布尔）	如果设置为 true，则表示不需要设置该参数；无需为该参数指定值，UI 即允许工作流执行。

如果某个参数为可选参数并且未设置，则其在 parameters 对象中传递给 execute 函数的成员将设置为 undefined。

工作流错误处理

如果在执行工作流期间发生错误，则会引发异常。如果异常未由工作流自身所捕获（或者如果工作流引发的异常未以其他方式捕获），则工作流将失败，并向用户显示有关异常的信息。要正确处理错误，应该捕获并处理异常。例如，在前面的示例中，尝试在不存在的项目中创建共享资源会导致未捕获的异常。此示例可以修改以捕获违例错误，并在项目不存在的情况下创建该项目：

```

var workflow = {
  name: 'Create share',
  description: 'Creates a new share in a business unit',
  parameters: {
    name: {
      label: 'Name of new share',
      type: 'String'
    },
    unit: {
      label: 'Business unit',
      type: 'ChooseOne',
      options: [ 'development', 'finance', 'qa', 'sales' ],
      optionlabels: [ 'Development', 'Finance',
        'Quality Assurance', 'Sales/Administrative' ],
    }
  },
  execute: function (params) {
    try {
      run('shares select ' + params.unit);
    } catch (err) {
      if (err.code !== EAKSH_ENTITY_BADSELECT)
        throw (err);

      /*
       * We haven't yet created a project that corresponds to
       * this business unit; create it now.
       */
      run('shares project ' + params.unit);
      run('commit');
      run('shares select ' + params.unit);
    }
  }
}
```

```
run('filesystem ' + params.name);
run('commit');
return ('Created new share "' + params.name + '"');
}
};
```

工作流输入验证

工作流可以选择通过添加 `validate` 成员来验证其输入，该成员将包含工作流参数作为成员的对象作为参数。`validate` 函数应返回一个对象，该对象的每个成员都使用未通过验证的参数命名，每个成员的值都是要向用户显示的验证错误消息。要将我们的示例扩展为在用户尝试创建已有共享资源时给出明确的错误：

```
var workflow = {
  name: 'Create share',
  description: 'Creates a new share in a business unit',
  parameters: {
    name: {
      label: 'Name of new share',
      type: 'String'
    },
    unit: {
      label: 'Business unit',
      type: 'ChooseOne',
      options: [ 'development', 'finance', 'qa', 'sales' ],
      optionlabels: [ 'Development', 'Finance',
        'Quality Assurance', 'Sales/Administrative' ],
    }
  },
  validate: function (params) {
    try {
      run('shares select ' + params.unit);
      run('select ' + params.name);
    } catch (err) {
      if (err.code == EAKSH_ENTITY_BADSELECT)
        return;
    }

    return ({ name: 'share already exists' });
  },
  execute: function (params) {
    try {
      run('shares select ' + params.unit);
    } catch (err) {
      if (err.code != EAKSH_ENTITY_BADSELECT)
        throw (err);
    }

    /*
     * We haven't yet created a project that corresponds to
     * this business unit; create it now.
     */
    run('shares project ' + params.unit);
    set('mountpoint', '/export/' + params.unit);
  }
};
```

```

        run('commit');
        run('shares select ' + params.unit);
    }

    run('filesystem ' + params.name);
    run('commit');
    return ('Created new share "' + params.name + '"');
}
};

```

工作流执行审计

工作流可以通过调用 `audit` 函数创建审计记录。`audit` 函数唯一的参数是要放置到审计日志中的字符串。

工作流执行报告

对于执行时间较长的复杂工作流，向执行工作流的用户提供明确的进度信息可能比较有用。要允许通过这种方式报告工作流的执行情况，`execute` 成员应该返回步骤数组。每个数组元素必须包含以下成员：

表 16-6 执行报告的必要成员

必要成员	类型	说明
<code>step</code>	字符串	表示执行步骤名称的字符串
<code>execute</code>	函数	执行工作流步骤的函数

与整个工作流上的 `execute` 函数一样，每个步骤的 `execute` 成员都将包含工作流参数的对象作为参数。例如，下面的工作流通过三个步骤创建新的项目、共享资源和审计记录：

```

var steps = [ {
    step: 'Checking for associated project',
    execute: function (params) {
        try {
            run('shares select ' + params.unit);
        } catch (err) {
            if (err.code !== EAKSH_ENTITY_BADSELECT)
                throw (err);
        }

        /*
         * We haven't yet created a project that corresponds to
         * this business unit; create it now.

```

```

        */
        run('shares project ' + params.unit);
        set('mountpoint', '/export/' + params.unit);
        run('commit');
        run('shares select ' + params.unit);
    }
}
}, {
    step: 'Creating share',
    execute: function (params) {
        run('filesystem ' + params.name);
        run('commit');
    }
}, {
    step: 'Creating audit record',
    execute: function (params) {
        audit('created "' + params.name + '" in "' + params.unit);
    }
}
];

var workflow = {
    name: 'Create share',
    description: 'Creates a new share in a business unit',
    parameters: {
        name: {
            label: 'Name of new share',
            type: 'String'
        },
        unit: {
            label: 'Business unit',
            type: 'ChooseOne',
            options: [ 'development', 'finance', 'qa', 'sales' ],
            optionlabels: [ 'Development', 'Finance',
                'Quality Assurance', 'Sales/Administrative' ],
        }
    },
    validate: function (params) {
        try {
            run('shares select ' + params.unit);
            run('select ' + params.name);
        } catch (err) {
            if (err.code == EAKSH_ENTITY_BADSELECT)
                return;
        }

        return ({ name: 'share already exists' });
    },
    execute: function (params) { return (steps); }
};

```

版本控制

与工作流相关的版本控制有两个方面：第一个是工作流依赖的 ZFSSA 软件版本表示，第二个是工作流自身的版本表示。版本控制通过两个可选的工作流成员进行表示：

表 16-7 版本控制的可选成员

可选成员	类型	说明
required	字符串	运行此工作流所需的 ZFSSA 软件的最早版本，包括最早年、月、日、内部版本和分支。
version	字符串	此工作流的版本，采用点分十进制 (major.minor.micro) 格式。

设备版本控制

要表示最低所需的设备软件版本，请将可选 `required` 字段添加到工作流。ZFSSA 的版本通过构建软件的年、月、日，后跟内部版本号以及分支号来标明，表示为 "year.month.day.build-branch"。例如 "2009.04.10.12-0" 是最初于 2009 年 4 月 10 日构建的软件的第十二个内部版本。要获取当前 ZFSSA 套件软件的版本，请运行 "configuration version get version" CLI 命令，或查看 BUI 中《[Oracle ZFS Storage Appliance 客户服务手册](#)》中的“系统”的 "Version"（版本）字段。下面是使用 `required` 字段的示例：

```
var workflow = {
  name: 'Configure FC',
  description: 'Configures fibre channel target groups',
  required: '2009.12.25,1-0',
  ...
}
```

如果工作流所需的软件版本高于 ZFSSA 上加载的版本，则尝试上载该工作流将失败，并显示一条消息指出不匹配。

工作流版本控制

除了指定所需的 ZFSSA 软件版本，工作流自身的版本也可以使用 `version` 字段来标明。此字符串表示工作流版本的主版本号、次版本号和微版本号，并允许计算机上存在同一工作流的多个版本。当上载工作流时，将删除同一工作流的所有兼容的早期版本。如果工作流具有相同的主版本号，则认为兼容，如果工作流具有更低的版本号，则认为早期版本。因此，上载版本为 "2.1" 的工作流将删除同一工作流的 "2.0" 版本（或 "2.0.1" 版本），但不会删除 "1.2" 版本或 "0.1" 版本。

作为警报操作的工作流

可选择性地执行工作流作为[第 9 章 警报配置](#)。要使工作流符合作为警报操作的资格，其 `alert` 操作必须设置为 `true`。

警报操作执行上下文

当作为警报操作执行时，工作流以创建工作流的用户身份运行。出于此原因，任何有资格作为警报操作的工作流必须将 `setid` 设置为 `true`。警报操作具有一个对象参数，该参数具有以下成员：

表 16-8 警报执行上下文的必要成员

必要成员	类型	说明
<code>class</code>	字符串	警报的类别。
<code>code</code>	字符串	警报的代码。
<code>items</code>	对象	说明警报的对象。
<code>timestamp</code>	日期	警报的时间。

`parameters` 对象的 `items` 成员具有以下成员：

表 16-9 项目成员的必要成员

必要成员	类型	说明
<code>url</code>	字符串	说明警报的网页 URL
<code>action</code>	字符串	用户响应警报应执行的操作。
<code>impact</code>	字符串	引发警报的事件的影响。
<code>description</code>	字符串	说明警报的用户可读的字符串。
<code>severity</code>	字符串	引发警报的事件的严重性。

审计警报操作

作为警报操作执行的工作流可以使用 `audit` 函数生成审计日志条目。建议通过 `audit` 函数将所有相关的调试信息生成到审计日志中。例如，下面是如果处于群集状态则执行故障转移的工作流 – 但是它审计所有无法重新引导故障：

```
var workflow = {
  name: 'Failover',
  description: 'Fail the node over to its clustered peer',
  alert: true,
  setid: true,
  execute: function (params) {
    /*
     * To failover, we first confirm that clustering is configured
```

```

* and that we are in the clustered state. We then reboot,
* which will force our peer to takeover. Note that we're
* being very conservative by only rebooting if in the
* AKCS_CLUSTERED state: there are other states in which it
* may well be valid to failback (e.g., we are in AKCS_OWNER,
* and our peer is AKCS_STRIPPED), but those states may also
* indicate aberrant operation, and we therefore refuse to
* failback. (Even in an active/passive clustered config, a
* FAILBACK should always be performed to transition the
* cluster peers from OWNER/STRIPPED to CLUSTERED/CLUSTERED.)
*/
var uuid = params.uuid;
var clustered = 'AKCS_CLUSTERED';

audit('attempting failover in response to alert ' + uuid);

try {
    run('configuration cluster');
} catch (err) {
    audit('could not get clustered state; aborting');
    return;
}

if ((state = get('state')) != clustered) {
    audit('state is ' + state + '; aborting');
    return;
}

if ((state = get('peer_state')) != clustered) {
    audit('peer state is ' + state + '; aborting');
    return;
}

run('cd /');
run('confirm maintenance system reboot');
}
};

```

使用调度的工作流

通过为工作流设置调度表，可以通过计时器事件启动工作流。必须将 `scheduled` 属性添加到工作流对象，并且该属性需要设置为 `true`。可以在将工作流加载到 ZFSSA 后通过 CLI 创建调度表，也可以将名为 `schedule` 的数组类型属性添加到对象工作流。

使用 CLI

将工作流加载到 ZFSSA 后，则可以通过 CLI 界面为其定义调度表，如下所示：

```

dory:> maintenance workflows
dory:maintenance workflows> "select workflow-002'"

```

```
dory:maintenance workflow-002> schedules
dory:maintenance workflow-002 schedules>create
dory:maintenance workflow-002 schedule (uncommitted)> set frequency=day
frequency = day (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> set hour=10
hour = 10 (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> set minute=05
minute = 05 (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> commit
dory:maintenance workflow-002 schedules> list
NAME          FREQUENCY      DAY          HH:MM
schedule-001   day            -            10:05
dory:maintenance workflow-002 schedules> create
dory:maintenance workflow-002 schedule (uncommitted)> set frequency=week
frequency = week (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> set day=Monday
day = Monday (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> set hour=13
hour = 13 (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> set minute=15
minute = 15 (uncommitted)
dory:maintenance workflow-002 schedule (uncommitted)> commit
dory:maintenance workflow-002 schedules> list
NAME          FREQUENCY      DAY          HH:MM
schedule-001   day            -            10:05
schedule-002   week           Monday       13:15
dory:maintenance workflow-002 schedules>
```

每个调度表条目都由以下属性组成：

表 16-10 调度属性

属性	类型	说明
NAME	字符串	系统生成的调度表名称
frequency	字符串	minute、halfhour、hour、day、week、month
day	字符串	指定具体哪天，可设置为：Monday、Tuesday、Wednesday、Thursday、Friday、或 Sunday。可以在 frequency 设置为 week 或 month 时设置
hour	字符串	00-23，指定调度表的小时部分，可以在 frequency 设置为 day、week 或 month 时指定。
minute	字符串	00-59，指定调度表的分钟部分。

对调度表进行编码

调度表还可以在工作流代码中指定为对象工作流的属性。这里使用的属性语法不同于 CLI 调度表创建。这里使用了三个属性：

表 16-11 调度属性

属性	类型	说明
offset	数字	确定定义期间的起点
period	数字	定义调度表的频率
unit	字符串	指定在 offset 和 period 定义中使用的单位是秒还是月

下面的代码示例说明了属性的用法。请注意，内嵌运算有助于让 offset 和 period 声明更加易读。

```
// Example of using Schedule definitions within a workflow
var MyTextObject = {
  MyVersion: '1.0',
  MyName: 'Example 9',
  MyDescription: 'Example of use of Timer',
  Origin: 'Oracle'
};
var MySchedules = [
  // half hr interval
  { offset: 0, period: 1800, units: "seconds" },
  // offset 2 days, 4hr, 30min , week interval
  {offset: 2*24*60*60+4*60*60+30*60, period: 604800,units: "seconds" }
];
var workflow = {
  name: MyTextObject.MyName,
  description: MyTextObject.MyDescription,
  version: MyTextObject.MyVersion,
  alert: false,
  setid: true,
  schedules: MySchedules,
  scheduled: true,
  origin: MyTextObject.Origin,
  execute: function () {
    audit('workflow started for timer; ');
  }
};
```

对象 MySchedules 中的属性 units 指定用于属性 offset 和 period 的单位类型。它们可以设置为 seconds 或 month。属性 period 指定事件的频率，offset 指定期间中的单位。在上面的示例中，第二个调度表的期间设置为一周，开始于第二天的 4:30。可以在 schedules 属性中定义多个调度表。

示例中的对象 MySchedules 使用以下三个属性：

- offset（偏移）– 这是调度表自 1970 年 1 月 1 日的起始偏移。属性 "units" 定义的单元中提供了此偏移量。
- period（周期）– 这是调度表重复之间的周期，该调度表也以通过属性“单位”定义的单位形式进行指定。
- units（单位）– 这可以定义为秒或月。

每周调度的起始时间是星期四。这是由于新时间定义为自 1970 年 1 月 1 日起，而这天是星期四。

在上例中，第二个调度表中的周期使用的起始偏移量为 2 天 + 4 小时 + 30 分钟。这导致开始日期变为 1970 年 1 月 3 日上午 4:30。此调度将无限地在每周六上午 4:30 执行。下文显示了 CLI 中的调度表。

```
<small>dory:> maintenance workflows
dory:maintenance workflows> list
WORKFLOW      NAME                                OWNER SETID ORIGIN              VERSION
workflow-000  Configure for Oracle Solaris Cluster NFS root  false Oracle Corporation  1.0.0
workflow-001  Unconfigure Oracle Solaris Cluster NFS root  false Oracle Corporation  1.0.0
workflow-002  Configure for Oracle Enterprise Manager Monitoring root  false Sun Microsystems, Inc. 1.1
workflow-003  Unconfigure Oracle Enterprise Manager Monitoring root  false Sun Microsystems, Inc. 1.0</small>
```

dory:maintenance workflow-002 schedules>

NAME	FREQUENCY	DAY	HH:MM
schedule-000	halfhour	-	--:00
schedule-001	week	Saturday	04:30

示例：设备类型选择

下面的示例 workflow 根据指定的驱动器类型创建工作表：

```
var steps = [ {
  step: 'Checking for existing worksheet',
  execute: function (params) {
    /*
     * In this step, we're going to see if the worksheet that
     * we're going to create already exists. If the worksheet
     * already exists, we blow it away if the user has indicated
     * that they desire this behavior. Note that we store our
     * derived worksheet name with the parameters, even though
     * it is not a parameter per se; this is explicitly allowed,
     * and it allows us to build state in one step that is
     * processed in another without requiring additional global
     * variables.
     */
    params.worksheet = 'Drilling down on ' + params.type + ' disks';

    try {
      run('analytics worksheets select name="' +
        params.worksheet + '"');

      if (params.overwrite) {
        run('confirm destroy');
        return;
      }

      throw ('Worksheet called "' + params.worksheet +
        '" already exists!');
    }
  }
}
```

```

    } catch (err) {
      if (err.code !== EAKSH_ENTITY_BADSELECT)
        throw (err);
    }
  }, {
    step: 'Finding disks of specified type',
    execute: function (params) {
      /*
       * In this step, we will iterate over all chassis, and for
       * each chassis iterates over all disks in the chassis,
       * looking for disks that match the specified type.
       */
      var chassis, name, disks;
      var i, j;

      run('cd /');
      run('maintenance hardware');

      chassis = list();
      params.disks = [];

      for (i = 0; i < chassis.length; i++) {
        run('select ' + chassis[i]);

        name = get('name');
        run('select disk');
        disks = list();

        for (j = 0; j < disks.length; j++) {
          run('select ' + disks[j]);

          if (get('use') == params.type) {
            params.disks.push(name + '/' +
              get('label'));
          }

          run('cd ..');
        }

        run('cd ../../');
      }

      if (params.disks.length === 0)
        throw ('No ' + params.type + ' disks found');
      run('cd /');
    }
  }, {
    step: 'Creating worksheet',
    execute: function (params) {
      /*
       * In this step, we're ready to actually create the worksheet
       * itself: we have the disks of the specified type and
       * we know that we can create the worksheet. Note that we
       * create several datasets: first, I/O bytes broken down
       * by disk, with each disk of the specified type highlighted
       * as a drilldown. Then, we create a separate dataset for
       * each disk of the specified type. Finally, note that we

```

```
* aren't saving the datasets -- we'll let the user do that
* from the created worksheet if they so desire. (It would
* be straightforward to add a boolean parameter to this
* workflow that allows that last behavior to be optionally
* changed.)
*/
var disks = [], i;

run('analytics worksheets');
run('create '' + params.worksheet + '');
run('select name='' + params.worksheet + '');
run('dataset');
run('set name=io.bytes[disk]');

for (i = 0; i < params.disks.length; i++)
  disks.push('' + params.disks[i] + '');

run('set drilldowns=' + disks.join(',')');
run('commit');

for (i = 0; i < params.disks.length; i++) {
  run('dataset');
  run('set name="io.bytes[disk=' +
    params.disks[i] + ']"');
  run('commit');
}
}
} ];

var workflow = {
  name: 'Disk drilldown',
  description: 'Creates a worksheet that drills down on system, ' +
    'cache, or log devices',
  parameters: {
    type: {
      label: 'Create a new worksheet drilling down on',
      type: 'ChooseOne',
      options: [ 'cache', 'log', 'system' ],
      optionlabels: [ 'Cache', 'Log', 'System' ]
    },
    overwrite: {
      label: 'Overwrite the worksheet if it exists',
      type: 'Boolean'
    }
  },
  execute: function (params) { return (steps); }
};
```

BUI

可通过单击加号图标将工作流上载到 ZFSSA，通过单击指定工作流的行执行工作流。

图 16-1

Workflows 5 Total		
NAME ^	DESCRIPTION	VERSION
Clear locks	Clear locks held on behalf of an NFS client	1.0.0
Configure for Oracle Enterprise Manager Monitoring	Sets up environment to be monitored by Oracle Enterprise Manager	1.1
Configure for Oracle Solaris Cluster NFS	Sets up environment for Oracle Solaris Cluster NFS	1.0.0
Unconfigure Oracle Enterprise Manager Monitoring	Removes the artifacts from the appliance used by Oracle Enterprise Manager	1.0
Unconfigure Oracle Solaris Cluster NFS	Removes the artifacts from the appliance used by Oracle Solaris Cluster NFS	1.0.0

CLI

工作流在 CLI 的 maintenance workflows 部分进行处理。

下载工作流

工作流通过 download 命令下载到 ZFSSA，与《Oracle ZFS Storage Appliance 客户服务手册》中的“系统”类似：

```
dory:maintenance workflows> download
dory:maintenance workflows download (uncommitted)> get
    url = (unset)
    user = (unset)
    password = (unset)

您必须将 "url" 属性设置为工作流的有效 URL。此 URL 可以位于本地网络，
也可以位于 Internet 上。URL 可以为 HTTP（以 "http://" 开头）或 FTP（以
"ftp://" 开头）。如果需要用户验证，可以将其包含在 URL 中（例如，"ftp://
myusername:mypasswd@myserver/export/foo"），也可以将用户名和密码放在 URL
之外，并设置用户和密码属性。

dory:maintenance workflows download (uncommitted)> set url=
    ftp://foo/example1.akwf
    url = ftp://foo/example1.akwf
dory:maintenance workflows download (uncommitted)> set user=bmc
    user = bmc
dory:maintenance workflows download (uncommitted)> set password
Enter password:
    password = *****
dory:maintenance workflows download (uncommitted)> commit
Transferred 138 of 138 (100%) ... done
```

查看工作流

要列出工作流，请在 maintenance workflows 上下文中使用 list 命令：

```
<small>dory:maintenance workflows> list
WORKFLOW    NAME                                OWNER SETID ORIGIN                                VERSION
workflow-000 Configure for Oracle Solaris Cluster NFS root  false Oracle Corporation  1.0.0
workflow-001 Unconfigure Oracle Solaris Cluster NFS root  false Oracle Corporation  1.0.0
workflow-002 Configure for Oracle Enterprise Manager Monitoring root  false Sun Microsystems, Inc.
1.1
workflow-003 Unconfigure Oracle Enterprise Manager Monitoring root  false Sun Microsystems, Inc.
1.0</small>
```

要查看工作流，请在 maintenance workflows 上下文中使用 show 命令：

```
dory:maintenance workflows> select workflow-001
dory:maintenance workflow-001> show
Properties:
    name = Configure for Oracle Solaris Cluster NFS
    description = Sets up environment for Oracle Solaris Cluster NFS
    owner = root
    origin = Oracle Corporation
    setid = false
    alert = false
    version = 1.0.0
    scheduled = false
```

要选择一个工作流，请使用 select 命令：

```
dory:maintenance workflows> select workflow-000
dory:maintenance workflow-000>
```

要获取工作流的属性，请从所选工作流的上下文中使用 get 命令：

```
dory:maintenance workflow-000> get
    name = Hello world
    description = Bids a greeting to the world
    owner = root
    origin = <local>
    setid = false
    alert = false
    scheduled = false
```

执行工作流

要执行工作流，请从所选工作流的上下文中使用 execute 命令。如果工作流不具有任何参数，则会直接执行：

```
dory:maintenance workflow-000> execute
hello world!
```

如果工作流具有参数，上下文将变为自持上下文，在其中必须指定参数：

```
dory:maintenance workflow-000> execute
dory:maintenance workflow-000 execute (uncommitted)> get
      type = (unset)
      overwrite = (unset)
```

在未设置必要参数前任何提交 workflow 执行的尝试都将导致显式错误的发生：

```
dory:maintenance workflow-000 execute (uncommitted)> commit
error: cannot execute workflow without setting property "type"
```

要执行 workflow，请设置指定的参数，然后使用 `commit` 命令：

```
dory:maintenance workflow-000 execute (uncommitted)> set type=system
      type = system
dory:maintenance workflow-000 execute (uncommitted)> set overwrite=true
      overwrite = true
dory:maintenance workflow-000 execute (uncommitted)> commit
```

如果 workflow 具有指定的步骤，这些步骤将通过 CLI 显示，例如：

```
dory:maintenance workflow-000 execute (uncommitted)> commit
Checking for existing worksheet ... done
Finding disks of specified type ... done
Creating worksheet ... done
```


集成

Oracle ZFS Storage Appliance 提供了一套完整的数据协议，可与各种应用程序主机进行通信。要提高应用程序性能或实现与您的应用环境更紧密集成，请遵循 "NAS Storage Documentation" (NAS 存储文档) 页面上的 "White Papers and Solutions Briefs" (白皮书和解决方案简介) 中概述的最佳实践。

- [“Symantec DMP/Storage Foundation”](#)

对于某些应用程序，在应用程序主机上安装软件可增强互操作性。以下文章概述了软件集成如何为存储管理员提供更好的体验。每个下载链接都包含打包的完整文档。

- [“用于 Oracle Solaris Cluster Geographic Edition 的 Sun ZFS Storage Appliance 插件” \[429\]](#)
- [“适用于 Oracle Solaris Cluster 的 Sun ZFS Storage Appliance 网络文件系统插件” \[429\]](#)
- [“适用于卷影副本服务软件的 Sun ZFS Storage Appliance 提供程序” \[431\]](#)
- [“适用于 Oracle Enterprise Manager Grid Control 的 Sun ZFS Storage 管理插件” \[429\]](#)
- [“适用于 Sun ZFS Storage Appliance 的 Oracle 虚拟机存储连接插件” \[431\]](#)
- [“适用于 VMware Site Recovery Manager 的 Sun ZFS Storage 7000 Storage Replication Adapter” \[433\]](#)
- [“ Oracle Intelligent Storage Protocol ” \[428\]](#)

您的设备也经过精心设计，可与其他 Oracle 产品进行无缝集成。例如，以下部分介绍如何将 ZFS Storage Appliance 配置为 Oracle Exadata Database Machine 和 Oracle SPARC SuperCluster 的备份目标。

- [“Oracle Exadata Database Machine 备份” \[398\]](#)
- [“为 ZFS Storage Appliance 备份配置 Oracle SPARC SuperCluster” \[424\]](#)

有关更多信息，请访问 "NAS Storage Documentation" (NAS 存储文档) 页面。

Oracle Exadata Database Machine 备份

在具有本地 QDR InfiniBand 和 10Gb 以太网连接选项的情况下，ZFS Storage Appliance 是可靠地备份 Oracle Exadata 的理想设备。可使用命令行工具部署 Oracle Exadata Backup Configuration Utility，或按照以下各节中的说明手动配置您的设备：

- [“手动配置 Sun ZFS Storage Appliance” \[398\]](#)
- [“针对 Sun ZFS Storage Appliance 配置 Oracle Exadata” \[401\]](#)

实用程序随附了内容全面的文档，其中包括如何从 Oracle Exadata 执行备份的说明。无论是以手动方式还是使用实用程序，采用任一方式都还需要在设备上配置网络连接和存储池。

有关使用 ZFS Storage Appliance 作为 Oracle Exadata 的备份目标的详细信息，请参见“NAS Storage Documentation”（NAS 存储文档）页面上的《Protecting Oracle Exadata with the Sun ZFS Storage Appliance: Configuration Best Practices》（《使用 Sun ZFS Storage Appliance 保护 Oracle Exadata：配置最佳实践》）白皮书。还提供了一个 Oracle ZFS Storage ZS3-4 群集，该群集作为 Oracle ZFS Storage ZS3-BA 随磁盘机框预先安装在机架上，可最大限度地降低设置的复杂性。将此设备与 Oracle Exadata 集成的过程与上文介绍的过程相同。

手动配置 Sun ZFS Storage Appliance

本节介绍了手动配置 ZFS Storage Appliance 以与 Oracle Exadata 结合使用的一般原则。有关详细信息，请参见“NAS Storage Documentation”（NAS 存储文档）页面上的《Protecting Oracle Exadata with the Sun ZFS Storage Appliance: Configuration Best Practices》（《使用 Sun ZFS Storage Appliance 保护 Oracle Exadata：配置最佳实践》）白皮书。

配置网络、池和共享资源

以下各节概述了优化 ZFS Storage Appliance 网络、存储池和共享资源配置以支持备份和恢复处理的最佳做法。

网络配置

本节介绍了如何配置 IP 网络多路径 (IP Network Multipathing, IPMP) 组以及如何在 ZFS Storage Appliance 中配置路由。

注意：如果您使用了 Oracle Exadata Backup Configuration Utility，请按照本节所述配置网络。有关详细信息，请查看“Best Practices”（最佳实践）白皮书。

对于需要其他 IB 连接的客户，还可以安装和配置其他 IB HCA (Host Channel Adapter, 主机通道适配器)。有关详细信息，请参见《Oracle ZFS Storage Appliance 安装指南》。

可将本节中介绍的原则应用到 10Gb 以太网实施，方法是将网络配置应用到 ixgbe 接口（而不是 ibp 接口）。10Gb 以太网实施可配置为主动/主动 IPMP。如果 ZFS Storage Appliance 位于 Oracle Exadata 以外的其他子网上，则可能需要创建从 ZFS Storage Appliance 到 Oracle Exadata 的静态路由。有关详细信息，请咨询网络管理员。

▼ 基本网络配置

1. 确保 ZFS Storage Appliance 已连接到 Oracle Exadata。
2. 为 `ibp0`、`ibp1`、`ibp2` 和 `ibp3` 配置地址 `0.0.0.0/8` (IPMP 要求如此)、连接模式和分区键 `ffff`。要确定 Oracle Exadata 系统使用的分区键，请以 root 用户身份运行以下命令：
`cat /sys/class/net/ib0/pkey`
3. 在 `ibd0` 和 `ibd3` 上配置活动/备用 IPMP 组，`ibd0` 为活动，`ibd3` 为备用。
4. 在 `ibd1` 和 `ibd2` 上配置活动/备用 IPMP 组，`ibd1` 为活动，`ibd2` 为备用。
5. 启用自适应路由，确保同一子网上的多个 IP 地址属于同一个机头时会相应地对通信进行负载均衡。群集故障转移后会出现此种情况。

池配置

本节介绍了在根据数据保护和性能要求，为 Oracle Recovery Manager (RMAN) 备份和恢复操作确定最适合 ZFS Storage Appliance 的池配置时需要注意的设计事项。

注意：如果您使用了 Oracle Exadata Backup Configuration Utility，请按照本节所述配置池。有关详细信息，请查看 "Best Practices" (最佳实践) 白皮书。

系统规划人员应基于以下原则考虑池保护：

- 对常规和容量优化系统使用基于奇偶特性的保护：
- * RAID-Z 用于保护具有随机工作负荷的系统免受单个驱动器故障的影响。
- * RAID-Z2 用于保护仅具有流工作负荷的系统免受双驱动器故障的影响。
- 对以增量方式应用备份的高性能系统使用镜像。
- 根据性能要求配置池：
- * 为管理优化系统配置一个池。
- * 为性能优化系统配置两个池。应当使用每个托盘中一半数量的驱动器来配置双池系统。

- 配置日志设备保护：
- * 分散读写日志设备用于 RAID-Z 和镜像池配置。
- * 镜像日志设备用于 RAID-Z2 池配置。

注意：如果您使用了 Oracle Exadata Backup Configuration Utility，请转到下一主题：[“针对 Sun ZFS Storage Appliance 配置 Oracle Exadata” \[401\]](#)。

共享资源配置

对于常规工作负荷，ZFS Storage Appliance 共享资源的默认选项是一个良好的起点。可按如下方式为 Oracle RMAN 备份和恢复操作优化 ZFS Storage Appliance 共享资源：

- 创建一个项目来存储与单个数据库的备份和恢复相关的所有共享资源。对于双池实施，请创建两个项目，每个池一个。
- 使用以下值配置支持 Oracle RMAN 备份和恢复工作负荷的共享资源：
- * Database record size (数据库记录大小) (recordsize) : 128kB
- * Synchronous write bias (同步写入偏向) (logbias) : "Throughput" (吞吐量) (用于处理备份集和映像副本) 或 "Latency" (延迟) (用于以增量方式应用的备份)
- * Cache device usage (高速缓存设备使用) (secondary cache) : "None" (无) (用于备份集) 或 "All" (全部) (支持以增量方式应用的备份或数据库克隆操作时)
- * Data compression (数据压缩) (compression) : "Off" (关闭) (用于性能优化的系统)、"LZJB" 或 "gzip-2" (用于容量优化的系统)
- * Number of shares per pool (每个池的共享资源数) : "1" (用于管理优化的系统)、"2" 或 "4" (用于性能优化的系统)

如果客户有相应要求，还可对用于支持 Oracle Exadata 备份和恢复的共享资源应用其他共享资源配置选项（例如较高级别的 gzip 压缩或复制）。

如果客户实施其他 ZFS Storage Appliance 数据服务，还应考虑特定于实施的测试，从而确定与之前介绍的做法的差异可能带来的影响。

配置 Oracle RMAN 和 Oracle 数据库实例

Oracle RMAN 是保护 Oracle Exadata 内容的基本组件。Oracle RMAN 可用于在 ZFS Storage Appliance 上创建 Oracle Exadata 内容的备份集、映像副本和以增量方式更新的备份。要优化从 Oracle Exadata 到 ZFS Storage Appliance 的 Oracle RMAN 备份的性能，数据库管理员应采用以下最佳做法：

- 对 Oracle RMAN 到数据库计算机的通道进行负载平衡，将负载平均分配到数据库计算机的各个节点上。

- 对 Oracle RMAN 到 ZFS Storage Appliance 的通道进行负载平衡，将负载平均分配到 Sun ZFS Storage Appliance 的各个共享资源和控制器上。

要优化 Oracle RMAN 通道到 ZFS Storage Appliance 的缓冲，可以调整几个隐藏实例参数的值。对于 Oracle Database 11g 发行版 2，可调整以下参数：

- 对于备份和恢复集：
 - *_backup_disk_bufcnt=64
 - *_backup_disk_bufsz=1048576
- 对于映像副本备份和恢复：
 - *_backup_file_bufcnt=64
 - *_backup_file_bufsz=1048576

有关调整这些参数和早期版本的 Oracle 数据库软件的等效参数的其他信息，请参见 ID 为 1072545.1 的文章：《*RMAN Performance Tuning Using Buffer Memory Parameters*》（《使用缓冲区内存参数调整 RMAN 性能》），网址为 [http://support.oracle.com \(http://support.oracle.com\)](http://support.oracle.com (http://support.oracle.com))。

Oracle Direct NFS (dNFS) 是为 Oracle RMAN 备份和恢复操作提供优异性能的高性能 NFS 客户机。应为需要最大限度提高备份和恢复操作吞吐量的客户配置 dNFS。

后续步骤

[“针对 Sun ZFS Storage Appliance 配置 Oracle Exadata” \[401\]](#)

针对 Sun ZFS Storage Appliance 配置 Oracle Exadata

本节包含一些样例脚本，用于说明如何将 ZFS Storage Appliance 连接到 Oracle Exadata。这些脚本用于在单池和双池 ZFS Storage Appliance 配置中支持名为 dbname 的数据库。

配置 Exadata：为 Sun ZFS Storage Appliance 配置 Oracle Exadata

大致实施步骤

实施步骤如下：

1. 设置在主机上挂载共享资源的目录结构（挂载点）。

2. 更新 `/etc/fstab` 以将从 ZFS Storage Appliance 导出的共享资源挂载到相应的挂载点。
3. 创建 `init.d` 服务以自动执行挂载和卸载共享资源过程。
4. 更新 `orantab` 文件以访问 ZFS Storage Appliance 导出的共享资源或在 `/etc/fstab` 中设置引导时挂载。
5. 在主机上挂载共享资源。
6. 更改挂载共享资源的权限以与 `ORACLE_HOME` 的权限设置匹配。
7. 重新启动 Oracle 数据库实例以获取 `orantab` 文件更改（可选）。

注意：如果您使用了 Oracle Exadata Backup Configuration Utility，则除了第 4 步和第 7 步以外的所有其他步骤均已由该实用程序代为执行。在下一节“具体实施步骤”中，您可选择执行“更新 `orantab` 以访问 ZFS Storage Appliance 导出”和“设置挂载共享资源的所有权”的第 2 步。

具体实施步骤

本节主题：

- 设置在主机上挂载共享资源的目录结构
- 更新 `/etc/fstab` 文件
- 创建 `init.d` 服务
- 更新 `orantab` 以访问 ZFS Storage Appliance 导出
- 在主机上挂载共享资源
- 设置挂载共享资源的所有权

设置在主机上挂载共享资源的目录结构

按如下所示设置共享资源在主机上的挂载点：

```
mkdir -p /zfssa/dbname/backup1
mkdir -p /zfssa/dbname/backup2
mkdir -p /zfssa/dbname/backup3
mkdir -p /zfssa/dbname/backup4
```

更新 `/etc/fstab` 文件

要更新 `/etc/fstab` 文件，请使用以下选项之一。

注意：下面列表中的 UNIX 换行转义符 (`\`) 表明单行代码在此处换到第二行。将有换行符的行输入 `fstab` 时，请删除 `\` 字符，然后两个行段合并到一行（用空格分隔）。

对于单池配置：

```
192.168.36.200:/export/dbname/backup1 /zfssa/dbname/backup1 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
192.168.36.200:/export/dbname/backup2 /zfssa/dbname/backup2 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
192.168.36.200:/export/dbname/backup3 /zfssa/dbname/backup3 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
192.168.36.200:/export/dbname/backup4 /zfssa/dbname/backup4 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
```

对于双池配置：

```
192.168.36.200:/export/dbname/backup1 /zfssa/dbname/backup1 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
192.168.36.201:/export/dbname/backup2 /zfssa/dbname/backup2 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
192.168.36.200:/export/dbname/backup3 /zfssa/dbname/backup3 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
192.168.36.201:/export/dbname/backup4 /zfssa/dbname/backup4 nfs \<br/>
rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,tcp,nfsvers= \<br/> 3,timeo=600 0 0
```

创建 init.d 服务

使用以下相应选项创建 init.d 服务。

```
# !/bin/sh
#
# zfssa_dbname: Mount ZFSSA project dbname for database dbname
#
# chkconfig: 345 61 19
# description: mounts ZFS Storage Appliance shares
#

start()
{
    mount /zfssa/dbname/backup1
    mount /zfssa/dbname/backup2
    mount /zfssa/dbname/backup3
    mount /zfssa/dbname/backup4
    echo "Starting $prog: "
}

stop()
{
    umount /zfssa/dbname/backup1
    umount /zfssa/dbname/backup2
    umount /zfssa/dbname/backup3
    umount /zfssa/dbname/backup4
    echo "Stopping $prog: "
```

```

}

case "$1" in
    start)
        start
        ;;
    stop)
        stop
        ;;
    restart)
        stop
        start
        ;;
    status)
        mount
        ;;
    *)
        echo "Usage: $0 {start|stop|restart|status}"
        exit 1
esac

```

(可选) 通过输入以下内容为 start-on-boot 启用 init.d 服务：

```
# chkconfig zfssa_dbname on
```

(可选) 使用 service 命令手动启动和停止服务：

```
# service zfssa_dbname start<br/># service zfssa_dbname stop
```

更新 oranfstab 以访问 ZFS Storage Appliance 导出

要更新 oranfstab 文件以访问 ZFS Storage Appliance 导出，请使用以下相应选项。

注意：如果您使用了 Oracle Exadata Backup Configuration Utility，则可以选择执行本过程。

对于单池配置：

```

server: 192.168.36.200
path: 192.168.36.200
export: /export/dbname/backup1 mount: /zfssa/dbname/backup1
export: /export/dbname/backup2 mount: /zfssa/dbname/backup2
export: /export/dbname/backup3 mount: /zfssa/dbname/backup3
export: /export/dbname/backup4 mount: /zfssa/dbname/backup4

```

对于双池配置：

```

server: 192.168.36.200
path: 192.168.36.200

```

```
export: /export/dbname/backup1 mount: /zfssa/dbname-2pool/backup1
export: /export/dbname/backup3 mount: /zfssa/dbname-2pool/backup3
server: 192.168.36.201
path: 192.168.36.201
export: /export/dbname/backup2 mount: /zfssa/dbname-2pool/backup2
export: /export/dbname/backup4 mount: /zfssa/dbname-2pool/backup4
```

在主机上挂载共享资源

要在主机上挂载共享资源，请输入以下两种选项之一：

```
# service mount_dbname start
```

或

```
# dcli -l root -g /home/oracle/dbs_group service mount_dbname start
```

设置挂载共享资源的所有权

更改挂载共享资源的权限设置以与 ORACLE_HOME 的权限设置匹配。在此示例中，用户和组的所有权均设置为 oracle:dba。

注意：如果您使用了 Oracle Exadata Backup Configuration Utility，则可以选择执行第 2 步，第 1 步已由该实用程序代为执行。

1. 输入以下两个选项之一：

```
# chown oracle:dba /zfssa/dbname/*
```

或

```
># dcli -l root -g /home/oracle/dbs_group chown oracle:dba/zfssa/dbname/*
```
2. 使用以下选项之一来重新启动 Oracle 数据库实例，从而获取对 oranfstab 文件所做的更改：

- 一次重新启动一个实例（滚动升级），例如：
- `$ srvctl stop instance -d dbname -i dbname1`
- `$ srvctl start instance -d dbname -i dbname1`
- `$ srvctl stop instance -d dbname -i dbname2`
- `$ srvctl start instance -d dbname -i dbname2`
- `$ srvctl stop instance -d dbname -i dbname3`
- `$ srvctl start instance -d dbname -i dbname3`
- `$ srvctl stop instance -d dbname -i dbname4`
- `$ srvctl start instance -d dbname -i dbname4`
- `$ srvctl stop instance -d dbname -i dbname5`
- `$ srvctl start instance -d dbname -i dbname5`

- `$ srvctl stop instance -d dbname -i dbname6`
- `$ srvctl start instance -d dbname -i dbname6`
- `$ srvctl stop instance -d dbname -i dbname7`
- `$ srvctl start instance -d dbname -i dbname7`
- `$ srvctl stop instance -d dbname -i dbname8`
- `$ srvctl start instance -d dbname -i dbname8`
- 重新启动整个数据库，例如：
- `$ srvctl stop database -d dbname`
- `$ srvctl start database -d dbname`

Oracle SPARC SuperCluster 备份

在具有本地 QDR InfiniBand 和 10Gb 以太网连接选项的情况下，ZFS Storage Appliance 是可靠地备份 Oracle SPARC SuperCluster 的理想设备。按照以下各节中的说明配置系统：

- [“配置 ZFS Storage Appliance 以便备份”](#)
- [“配置 Oracle SPARC SuperCluster 以进行 ZFS Storage Appliance 备份”](#)

有关使用 ZFS Storage Appliance 作为 Oracle SPARC SuperCluster 的备份目标的详细信息，请参见“NAS Storage Documentation”（NAS 存储文档）页面上的《Configuring a Sun ZFS Backup Appliance with Oracle SPARC SuperCluster》（《配置 Oracle SPARC SuperCluster 和 Sun ZFS Backup Appliance》）白皮书。还提供了—个 Oracle ZFS Storage ZS3-4 群集，该群集作为 Oracle ZFS Storage ZS3-BA 随磁盘机框预先安装在机架上，可最大限度地降低设置的复杂性。将此设备与 Oracle SPARC SuperCluster 集成的过程与上文介绍的过程相同。

配置 ZFS Storage Appliance 以便备份

本节介绍了配置 ZFS Storage Appliance 以与 Oracle SPARC SuperCluster 结合使用进行备份的一般原则。有关详细信息，请参见“NAS Storage Documentation”（NAS 存储文档）页面上的《Configuring a Sun ZFS Backup Appliance with Oracle SPARC SuperCluster》（《配置 Oracle SPARC SuperCluster 和 Sun ZFS Backup Appliance》）白皮书。以下示例显示了带两个控制器（机头）和四个磁盘机框的 ZFS Storage Appliance。

本节主题：

- [“配置 ZFS Storage Appliance InfiniBand 数据链路” \[407\]](#)

- “配置 Oracle SPARC SuperCluster InfiniBand 交换机以添加 ZFS Storage Appliance” [408]
- “为单个 IP 连接配置 ZFS Storage Appliance 网络” [410]
- “配置 ZFS Storage Appliance 网络以实现主动-主动配置” [411]
- “配置 ZFS Storage Appliance 存储池” [413]
- “配置 ZFS Storage Appliance 共享资源” [413]
- “配置 ZFS Storage Appliance DTrace Analytics” [414]
- “配置客户机 NFS 挂载” [415]
- “调整 Solaris 11 网络和内核” [415]
- “配置 Oracle Direct NFS (dNFS)” [416]
- “为 Oracle RMAN 备份和恢复操作调整 Oracle 数据库实例” [417]
- “Creating Dedicated Services for Oracle RMAN Operations”“为 Oracle RMAN 操作创建专用服务” [419]
- “配置 Oracle RMAN” [419]

配置 ZFS Storage Appliance InfiniBand 数据链路

按照本节中的步骤可配置各个 ZFS Storage Appliance InfiniBand 连接。可使用在此过程中为 InfiniBand HBA 端口记录的八个 GUID 来配置下一个过程中的 Oracle SPARC SuperCluster InfiniBand 交换机。

1. 将 ZFS Storage Appliance 连接到 Oracle SPARC SuperCluster，如“NAS 存储文档”页面上的《Configuring a Sun ZFS Backup Appliance with Oracle SPARC SuperCluster》（《配置 Oracle SPARC SuperCluster 和 Sun ZFS Backup Appliance》）白皮书中所述。
2. 登录到机头 1 的浏览器用户界面 (Browser User Interface, BUI)，并导航到 "Configuration"（配置）> "Network"（网络）。
3. 单击 "Datalinks"（数据链路）旁的加号图标。将打开 "Network Datalink"（网络数据链路）对话框。
4. 如下所述完成该对话框。
 - 选中 IB Partition 框。
 - 为此数据链路输入一个有意义的名称。
 - 将 Partition Key 设置为 8503。
 - 为链路模式选择连接模式。
 - 切勿选中 LACP Aggregation 框。
 - 选择 Partition Device ibp0。
 - 记录 GUID 编号（例如，21280001ef43bb）并单击 "Apply"（应用）。

5. 针对剩余的每个 InfiniBand 接口 (ibp1、ibp2 和 ibp3) 重复步骤 3 和 4。
6. 针对机头 2 重复步骤 2 到步骤 5。

配置 Oracle SPARC SuperCluster InfiniBand 交换机以添加 ZFS Storage Appliance

在此过程中，ZFS Storage Appliance Infiniband HBA 的 GUID 将添加到现有的 Oracle SPARC SuperCluster InfiniBand 配置中。通过添加这些端口并使用分区键 8503，可实现这两个设备之间的通信。

1. 以 root 身份登录 Oracle SPARC SuperCluster InfiniBand 中心交换机。默认情况下，将为此中心交换机提供一个主机名 <sscid>sw-ib1，其中，<sscid> 是提供给整个 Oracle SPARC SuperCluster 系统的前缀名。在下例中，<sscid> 是 aiessc。

```
login as: root
root@aiesscsw-ib1's password:
Last login: Tue Sep 25 08:19:01 2013 from dhcp-brm-bl5-204-3e
east-10-135-75-254.usdhcp.oraclecorp.com
```

2. 输入命令 enablesm 以确认此交换机正在运行子网管理器（如未运行，则此命令将启动子网管理器）。


```
[root@aiesscsw-ib1 ~]# enablesm
opensm (pid 15906) is already running...
Starting partitiond daemon
/usr/local/util/partitiond is already running
(You may also perform a 'restart' if wanted)
```

3. 输入命令 getmaster 以确认这是配置的主交换机。如果此主交换机未在中心交换机上运行，注销后再次登录指定主交换机，以处理此过程的剩余部分。


```
[root@aiesscsw-ib1 ~]# getmaster
Local SM enabled and running
20130913 10:16:51 Master SubnetManager on sm lid 13 sm guid
0x2128e8ac27a0a0 : SUN DCS 36P QDR aiesscsw-ib1.us.oracle.com
[root@aiesscsw-ib1 ~]#
```

4. 根据记录的备份过程备份交换机配置 (http://docs.oracle.com/cd/E26698_01/index.html (http://docs.oracle.com/cd/E26698_01/index.html))。

5. 输入命令 smpartition list active 以确认分区键 0x0503 已分配给分区名称 "sto" (sto = 0x0503)。
分区键在 ZFS Storage Appliance 数据链路上设置为 8503，但 InfiniBand 交换机报告为 0503。这是有意如此，因为 InfiniBand 协议保留十六进制分区键 (pkey) 的最高有效位 (0x8000) 供自己使用。因此，pkey 0x8503 和 0x0503 是相同的。


```
[root@aiesscsw-ib1 ~]# smpartition list active
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 11
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=
full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
ic1s10 = 0x0501, ipoib, defmember=full:
0x0021280001ef30f7,
0x0021280001ef33bf,
0x0021280001ef30b7,
0x0021280001ef314b;
ic2s10 = 0x0502, ipoib, defmember=full:
0x0021280001ef30f8,
0x0021280001ef33c0,
0x0021280001ef30b8,
0x0021280001ef314c;
sto = 0x0503, ipoib, defmember=full:
0x0021280001ef43f8,
0x0021280001ef43b7,
0x0021280001cf90c0,
0x0021280001ef43bb,
...more...
```

6. 将 ZFS Storage Appliance 添加到 InfiniBand 配置：

- 输入命令 `smpartition start` 以启动重新配置会话。


```
# smpartition start<br/>
[root@aiesscsw-ib1 ~]# smpartition start
```

- 输入命令 `smpartition add` 以将八个新 GUID 添加到配置中。


```
# smpartition add -n sto -port <GUID1> <GUID2> <GUID3> ... <GUID8><br/>
[root@aiesscsw-ib1 ~]# smpartition add -n sto -port
21280001ef43bb 21280001ef43bc 21280001cf90bf 21280001cf90c0
21280001ef43f7 21280001ef43f8 21280001ef43b7 21280001ef43b8
```

- 输入命令 `smpartition list modified` 以确认已正确添加了新 GUID。


```
# smpartition list modified<br/>
[root@aiesscsw-ib1 ~]# smpartition list modified
# Sun DCS IB partition config file
<nowiki># This file is generated, do not edit
#! version_number : 11
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=
full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
ic1s10 = 0x0501, ipoib, defmember=full:
0x0021280001ef30f7,
0x0021280001ef33bf,
0x0021280001ef30b7,
```

```

0x0021280001ef314b;
ic2s10 = 0x0502,ipoib,defmember=full:
0x0021280001ef30f8,
0x0021280001ef33c0,
0x0021280001ef30b8,
0x0021280001ef314c;
sto = 0x0503,ipoib,defmember=full:
0x0021280001ef43f8,
0x0021280001ef43b7,
0x0021280001cf90c0,
0x0021280001ef43bb,
0x0021280001ef43bc,
0x0021280001cf90bf,
0x0021280001ef43b8,
0x0021280001ef43f7,
0x0021280001ef3048,
0x0021280001ef30af,
0x0021280001ef30f8,
0x0021280001ef30f7,
0x0021280001ef33c0,
0x0021280001ef33bf,
0x0021280001ef30cc,
0x0021280001ef342b,
0x0021280001ef30b8,
0x0021280001ef30b7,
0x0021280001ef314c,
0x0021280001ef314b,
0x0021280001efec65,
0x0021280001efec66,
0x0021280001efecb1,
0x0021280001efecb2;

```

- 输入命令 `smpartition commit` 以应用新配置并将配置更改传播到配置中的所有 InfiniBand 交换机。


```

# smpartition commit<br/>
[root@aiescsw-ib1 ~]# smpartition commit
[root@aiescsw-ib1 ~]#

```

7. 注销 InfiniBand 交换机。

8. 根据记录的备份过程备份 InfiniBand 配置 (http://docs.oracle.com/cd/E26698_01/index.html (http://docs.oracle.com/cd/E26698_01/index.html))。

为单个 IP 连接配置 ZFS Storage Appliance 网络

此配置仅适用无外部叶交换机的 Oracle SPARC SuperCluster T5。为了实现最佳的故障转移和性能，所有其他配置都可使用主动-主动配置（下一节）。

通过使用以下过程为端口 1 配置所需的 IP 地址，从而为 ZFS Storage Appliance InfiniBand 端口配置网络连接和简单群集故障转移。

1. 登录到机头 1 的 BUI，并导航到 "Configuration"（配置）> "Network"（网络）。
2. 单击 "Interfaces"（接口）旁的加号图标。将打开 "Network Interface"（网络接口）对话框。
3. 如下所述完成该对话框。
 - 为网络接口输入一个有意义的名称。
 - 确认已选中 Enable Interface（启用接口）。
 - 确认已选中 Allow Administration（允许管理）。
 - 确认已选中 Use IPv4 Protocol（使用 IPv4 协议）。
 - 确认 Configure with（配置方式）菜单选择为 Static Address List（静态地址列表）。
 - 在下框中输入所需的 IP 地址和相应的网络掩码。
 - 确认未选中 Use IPv6 Protocol（使用 IPv6 协议）。
 - 为 ibp0 选择数据链路，并单击 "Apply"（应用）。
4. 使用 ibp2 作为数据链路，在机头 2 上重复步骤 1 到 3。

配置 ZFS Storage Appliance 网络以实现主动-主动配置

在 ZFS Storage Appliance 上配置 InfiniBand 端口，以实现 IP 多路径。在专用存储子网上，每个 ZFS Storage Appliance 机头都需要四个 IP 地址（因此共需八个地址），因为接口将在主动-主动配置下运行。

1. 将各个 InfiniBand 数据链路配置为其自己的网络接口。
 - 登录到机头 1 的 BUI，并导航到 "Configuration"（配置）> "Network"（网络）。
 - 单击 "Interfaces"（接口）旁的加号图标。将打开 "Network Interface"（网络接口）对话框。
 - 如下所述完成该对话框。
 - * 为网络接口输入一个有意义的名称。
 - * 确认已选中 Enable Interface（启用接口）。
 - * 确认已选中 Allow Administration（允许管理）。
 - * 确认已选中 Use IPv4 Protocol（使用 IPv4 协议）。
 - * 确认 "Configure with"（配置方式）菜单选项为 "Static Address List"（静态地址列表）。
 - * 在下框中，输入 0.0.0.0/8。
 - * 确认未选中 Use IPv6 Protocol（使用 IPv6 协议）。

- * 为 ibp0 选择数据链路，并单击 "Apply"（应用）。
- 针对剩余的数据链路（ibp1、ibp2 和 ibp3）重复第二和第三个子步骤。
- 在机头 2 上重复第一到第四个子步骤。

2. 在机头 1 上配置 IPMP 接口。

- 登录到机头 1 的 BUI，并导航到 "Configuration"（配置）> "Network"（网络）。
- 单击 "Interfaces"（接口）旁的加号图标。将打开 "Network Interface"（网络接口）对话框。
- 如下所述完成该对话框。
- * 为 IPMP 网络接口输入一个有意义的名称。
- * 确认已选中 Enable Interface（启用接口）。
- * 确认已选中 Allow Administration（允许管理）。
- * 确认已选中 Use IPv4 Protocol（使用 IPv4 协议）。
- * 确认 Configure with（配置方式）菜单选项为 Static Address List（静态地址列表）。
- * 单击空框旁的加号三次，将显示四个空框。
- * 在每个空框中输入一个为 InfiniBand 连接保留的 IP 地址以及为该地址指派的 /24 网络掩码。最佳做法是不使用块中的连续 IP 地址，而是隔一个用一个（例如，全部是奇数或偶数位的）。
- * 确认未选中 Use IPv6 Protocol（使用 IPv6 协议）。
- * 选中 IP MultiPathing Group（IP 多重路径组）框。
- * 选中与数据链路 ibp0 和 ibp3 相对应的接口旁边的框。
- * 确认两个接口都设置为 Active（活动），然后单击 "Apply"（应用）。
- 从 "Configuration"（配置）> "Network"（网络）中，单击 "Routing"（路由）。
- 单击对应于 Adaptive（自适应）的多宿主模式。

3. 在机头 2 上配置 IPMP 接口。

- 登录到机头 2 的 BUI，并导航到 "Configuration"（配置）> "Network"（网络）。
- 单击 "Interfaces"（接口）旁的加号图标。将打开 "Network Interface"（网络接口）对话框。
- 如下所述完成该对话框。
- * 为 IPMP 网络接口输入一个有意义的名称。
- * 确认已选中 Enable Interface（启用接口）。
- * 确认已选中 Allow Administration（允许管理）。
- * 确认已选中 Use IPv4 Protocol（使用 IPv4 协议）。
- * 确认 Configure with（配置方式）菜单选项为 Static Address List（静态地址列表）。
- * 单击空框旁的加号三次，将显示四个空框。

- * 在每个空框中输入为 InfiniBand 连接保留的四个 IP 地址之一以及为该地址指派的 /24 网络掩码。这些地址必须是机头 1 上未使用的地址。
- * 确认未选中 Use IPv6 Protocol (使用 IPv6 协议)。
- * 选中 IP MultiPathing Group (IP 多重路径组) 框。
- * 选中与数据链路 ibp1 和 ibp2 相对应的接口旁边的框。
- * 确认两个接口都设置为 Active (活动)，然后单击 "Apply" (应用)。
- 从 "Configuration" (配置) > "Network" (网络) 中，单击 "Routing" (路由)。
- 单击对应于 Adaptive (自适应) 的多宿主模式。

4. 确认与 Oracle SPARC SuperCluster 节点的连接。确认每个节点都能 ping 出 ZFS Storage Appliance 上各个 IPMP 组中使用的八个地址。将这些 IP 地址添加到各个节点的 /etc/inet/hosts 表中。

配置 ZFS Storage Appliance 存储池

池配置向逻辑存储池分配物理磁盘驱动器资源，以用于备份数据存储。为了最大程度提高系统吞吐量，可通过向每个存储池分配各个驱动器托盘的一半物理驱动器，从而配置两个大小相同的存储池。

当配置的两个池具有相同 RAID 保护配置文件时，ZFS Storage Appliance 管理软件会显示一条有关效率的警告消息。如果是为高性能 Oracle RMAN 备份解决方案而配置，可安全地忽略此消息。

配置 ZFS Storage Appliance 共享资源

共享资源配置是指设置和运行 NFS 挂载点以便客户机进行访问的过程。应为 Oracle SPARC SuperCluster 配置创建两个项目：每个池一个项目。项目是一个为一组共享资源提供高级管理接口点的实体。要优化共享资源管理，可更新项目中所包含共享资源的默认挂载点，使其引用数据库名称，例如 /export/dbname。对于性能优化的系统，为每个池中的各个项目创建四个共享资源，共八个共享资源（每个机头四个共享资源）。要配置项目，请执行以下操作：

1. 登录到机头 1 的 BUI，并导航到 "Shares" (共享资源) > "Projects" (项目)。
2. 单击 "Projects" (项目) 旁的加号，为此项目输入一个有意义的名称，然后单击 "Apply" (应用)。由于将在另一个机头上创建类似项目，为机头 1 的项目提供唯一的名称，例如 H1-mydb。
3. 单击新项目名称旁的铅笔图标，以编辑此项目。
4. 单击 "General" (常规)，并如下所述完成属性：

- 更改 Mountpoint 以包括数据库名称（例如，/export/H1-mysdb）。
- 将 Synchronous write bias（同步写入偏向）从 Latency（延迟）更改为 Throughput（吞吐量），然后单击 "Apply"（应用）。

5. 单击 "Protocols"（协议），并如下所述添加 NFS 例外情况：

- 单击 "NFS Exceptions"（NFS 例外情况）旁的加号。
- 将 Type（类型）更改为 Network（网络）。
- 输入 InfiniBand 网络的子网和网络掩码（例如，/24）。
- 将 Access Mode（访问模式）更改为 Read/Write（读/写）。
- 确认 Charset（字符集）设置为 default（默认）。
- 选中 Root Access（Root 访问权限）框并单击 "Apply"（应用）。

6. 单击 "General"（常规）旁的 "Shares"（共享资源）。

7. 为机头 1 创建四个文件系统，并为它们提供唯一的名称，使其与机头 2 的文件系统名称不同。要交错执行备份流以在两个机头上分配数据，从而提供更佳的性能，可让机头 1 使用奇数编号的名称，例如 backup1、backup3、backup5 和 backup7；让机头 2 使用偶数编号的名称，例如 backup2、backup4、backup6 和 backup8。要创建文件系统，单击 "Filesystems"（文件系统）旁的加号，输入文件系统 (backup1) 的名称，然后单击 "Apply"（应用）。重复此步骤以创建剩余的三个文件系统 (backup3、backup5 和 backup7)。

8. 针对机头 1 重复步骤 2 到步骤 7。记住要使用唯一的项目名称（例如，H2-mysdb），并为这些文件系统名称指定偶数编号的备份 ID (backup2、backup4、backup6 和 backup8)。

配置 ZFS Storage Appliance DTrace Analytics

ZFS Storage Appliance 包括一个名为 DTrace Analytics 的综合性能分析工具。DTrace Analytics 是一种用于监视重要子系统性能记账统计信息的框架。应监视部分可用记账统计信息，以提供有关 Oracle RMAN 备份和恢复工作负荷的有效性和性能的综合数据。

当在 ZFS Storage Appliance 上配置了高级分析时 ("Configuration"（配置） > "Preferences"（首选项） > "Enable Advanced Analytics"（启用高级 Analytics))，以下 Analytics 可用：

- CPU: Percent utilization broken down by CPU mode（CPU：按 CPU 模式细分的利用率百分比）
- Disk: Average number of I/O operations broken down by state of operation（磁盘：按操作状态细分的平均 I/O 操作数）
- Disk: I/O bytes per second broken down by type of operation（磁盘：按操作类型细分的每秒 I/O 字节数）

- Disk: I/O operations per second broken down by latency (磁盘：按延迟细分的每秒 I/O 操作数)
- Disk: Disks with utilization of at least 95 percent broken down by disk (磁盘：按磁盘细分的利用率至少为 95% 的磁盘)
- Network: Interface bytes per second broken down by direction (网络：按方向细分的每秒接口字节数)
- Network: Interface bytes per second broken down by interface (网络：按接口细分的每秒接口字节数)
- Protocol: NFSv3 operations per second broken down by size (协议：按大小细分的每秒 NFSv3 操作数)
- Protocol: NFSv3 operations per second broken down by type of operation (协议：按操作类型细分的每秒 NFSv3 操作数)
- Protocol: NFSv3 operations per second of type read broken down by latency (协议：按延迟细分的每秒读取类型的 NFSv3 操作数)
- Protocol: NFSv3 operations per second of type write broken down by latency (协议：按延迟细分的每秒写入类型的 NFSv3 操作数)
- Protocol: NFSv3 operations per second of type read broken down by size (协议：按大小细分的每秒读取类型的 NFSv3 操作数)
- Protocol: NFSv3 operations per second of type write broken down by size (协议：按大小细分的每秒写入类型的 NFSv3 操作数)

实施这些记账统计信息可帮助最终用户在一定程度上了解具体实施的资源消耗的即时和历史情况及服务质量 (quality of service, QoS)。

配置客户机 NFS 挂载

配置 ZFS Storage Appliance 时，任何访问此设备的服务器（包括 Oracle SPARC SuperCluster 节点）都会被视为客户机。客户机 NFS 挂载的配置过程包括创建用来访问 ZFS Storage Appliance 的目标目录结构，以及为实现最佳系统性能而必需的特定 NFS 挂载选项。Solaris 客户机的挂载选项为：

```
rw,bg,hard,nointr,rsz=1048576,wsz=1048576,proto=tcp,vers=3,forcedirectio
```

在 ZFS Storage Appliance 上创建的目录的挂载点也应在各个 Oracle SPARC SuperCluster 节点上创建，并添加到其 `/etc/inet/hosts` 表中。

调整 Solaris 11 网络和内核

应将以下条目添加到各个 Oracle SPARC SuperCluster 节点的 `/etc/system` 文件：

```
set rpcmod:clnt_max_conns = 8
set nfs:nfs3_bsize = 131072
```

此外，每次重新引导 Oracle SPARC SuperCluster 时，需要在各个 Oracle SPARC SuperCluster 节点上运行以下命令：

```
/usr/sbin/ndd -set /dev/tcp tcp_max_buf 2097152
/usr/sbin/ndd -set /dev/tcp tcp_xmit_hiwat 1048576
/usr/sbin/ndd -set /dev/tcp tcp_recv_hiwat 1048576
```

要实现最佳性能，可能需要进行其他调整。有关最新信息，请参阅《Oracle SPARC SuperCluster Tunables》（《Oracle SPARC SuperCluster 可调参数》）文档 1474401.1，网址为 <http://support.oracle.com> (<http://support.oracle.com>)。同时，2013 年 1 月的 QFSDP 版本添加了可自动设置可调参数的 "ssctuner" 工具。有关其他信息，请参阅 Oracle SPARC SuperCluster 发行说明。

配置 Oracle Direct NFS (dNFS)

在每个 Oracle SPARC SuperCluster 节点上，对 dNFS 执行如下配置：

1. 关闭正在运行的 Oracle 数据库软件实例。
2. 将目录更改为 \$ORACLE_HOME/rdbms/lib。
3. 启用 dNFS：

```
make -f $ORACLE_HOME/rdbms/lib/ins_rdbms.mk dnfs_on
```

4. 使用特定于配置的服务器名称、路径名称和导出名称更新 oranfstab 文件（位于 /\$ORACLE_HOME/dbs），其中，

- 服务器参数引用 ZFS Storage Appliance 机头在 InfiniBand 网络上的本地名称。
- 路径参数应反映配置期间指定的机头的地址。
- 导出参数应反映类似于在 /etc/vfstab 中创建的条目的挂载点。这些条目应类似于以下条目。

对于单 IP 配置（仅限无外部交换机的 Oracle SPARC SuperCluster T5）：

```
server: aie-zba-h1-stor
path: 192.168.30.100
export: /export/test1/backup1 mount: /zba/test1/backup1
export: /export/test1/backup3 mount: /zba/test1/backup3
export: /export/test1/backup5 mount: /zba/test1/backup5
export: /export/test1/backup7 mount: /zba/test1/backup7
server: aie-zba-h2-stor
```



```

path: 192.168.30.101
export: /export/test1/backup2 mount: /zba/test1/backup2
export: /export/test1/backup4 mount: /zba/test1/backup4
export: /export/test1/backup6 mount: /zba/test1/backup6
export: /export/test1/backup8 mount: /zba/test1/backup8<br/>

```

对于 IPMP 组配置（所有其他项目）：

```

server: aie-zba-h1-stor
path: 192.168.30.100
path: 192.168.30.102
path: 192.168.30.104
path: 192.168.30.106
export: /export/test1/backup1 mount: /zba/test1/backup1
export: /export/test1/backup3 mount: /zba/test1/backup3
export: /export/test1/backup5 mount: /zba/test1/backup5
export: /export/test1/backup7 mount: /zba/test1/backup7
server: aie-zba-h2-stor
path: 192.168.30.101
path: 192.168.30.103
path: 192.168.30.105
path: 192.168.30.107
export: /export/test1/backup2 mount: /zba/test1/backup2
export: /export/test1/backup4 mount: /zba/test1/backup4
export: /export/test1/backup6 mount: /zba/test1/backup6
export: /export/test1/backup8 mount: /zba/test1/backup8

```

5. 重新启动 Oracle 数据库软件实例。

为 Oracle RMAN 备份和恢复操作调整 Oracle 数据库实例

使用 Oracle RMAN 和 ZFS Storage Appliance 优化高带宽备份和恢复操作需要调整控制 I/O 缓冲的实例参数。有关如何调整这些参数的信息，请参见 ID 为 1072545.1 的文章：《RMAN Performance Tuning Using Buffer Memory Parameters》（《使用缓冲区内存参数调整 RMAN 性能》），网址为 <http://support.oracle.com> (<http://support.oracle.com>)。

对于 Oracle SPARC SuperCluster，应考虑调整四个参数：

- `_backup_disk_bufcnt` – 用于处理备份集的缓冲区数
- `_backup_disk_bufsz` – 用于处理备份集的缓冲区大小
- `_backup_file_bufcnt` – 用于处理映像副本的缓冲区数
- `_backup_file_bufsz` – 用于处理映像副本的缓冲区大小

对于针对备份集和映像副本的备份和恢复操作，将缓冲区数设置为 64，将缓冲区大小设置为 1 MB：

```
SQL> alter system set "_backup_disk_bufcnt"=64;
SQL> alter system set "_backup_file_bufcnt"=64;
SQL> alter system set "_backup_disk_bufsz"=1048576;
SQL> alter system set "_backup_file_bufsz"=1048576;
```

可将这些命令添加到 SPFILE 以永久性配置它们，或可在用于执行备份或恢复操作的 Oracle RMAN 运行块中动态地设置这些命令。

以下代码片段显示了如何针对备份和恢复操作动态地调整缓冲区大小和计数。

■ 备份集备份：

```
run
{<br/>
  sql 'alter system set "_backup_disk_bufcnt"=64';<br/>
  sql 'alter system set "_backup_disk_bufsz"=1048576';<br/>
  allocate channel...
...<br/>
  backup as backupset database;
}
```

■ 备份集恢复：

```
run
{<br/>
  sql 'alter system set "_backup_disk_bufcnt"=64';<br/>
  sql 'alter system set "_backup_disk_bufsz"=1048576';<br/>
  allocate channel...
...<br/>
  restore database;
}
```

■ 映像副本备份：

```
run
{<br/>
  sql 'alter system set "_backup_file_bufcnt"=64';<br/>
  sql 'alter system set "_backup_file_bufsz"=1048576';<br/>
  allocate channel...
...<br/>
  backup as copy database;
}
```

■ 映像副本恢复：

```
run
{<br/>
  sql 'alter system set "_backup_file_bufcnt"=64';<br/>
  sql 'alter system set "_backup_file_bufsz"=1048576';<br/>
  allocate channel...
...<br/>
  restore database;
}
```

```
}
```

执行以增量方式应用的备份操作需要读取增量备份集和写入到映像副本。要为以增量方式应用的备份调整缓冲区，请运行以下命令：

```
run
{
  sql 'alter system set "_backup_disk_bufcnt"=64';
  sql 'alter system set "_backup_disk_bufsz"=1048576';
  sql 'alter system set "_backup_file_bufcnt"=64';
  sql 'alter system set "_backup_file_bufsz"=1048576';
  allocate channel...
  ...
  recover copy of database;
}
```

为 Oracle RMAN 操作创建专用服务

可配置两个专用于 Oracle RMAN 处理操作的服务，以优化对负载平衡、高可用性和升级的管理。可对这些服务进行负载平衡，将这些服务平均分配到 Oracle SPARC SuperCluster 系统的所有节点上。可通过以下方法来优化可用性和性能：将这些服务配置为在首选实例上运行，同时将它们准备为将故障转移到群集内的任何其他实例上。如果配置了这些服务，升级四分之一或一半的机架 Oracle SPARC SuperCluster 系统无需更改 Oracle RMAN 运行块的连接字符串。

srvctl 实用程序可用于安装用于 Oracle RMAN 处理操作的服务。以下代码片段显示了如何创建两个平均分布在四节点群集上并设置为将故障转移到群集内任何其他节点的服务。在此例中，为名为 dbname 的数据库安装这些服务，且将这些服务命名为 dbname_bkup。

```
srvctl add service -d dbname -r dbname1 -a dbname2 -s dbname_bkup1
srvctl start service -d dbname -s dbname_bkup1
srvctl add service -d dbname -r dbname2 -a dbname1 -s dbname_bkup2
srvctl start service -d dbname -s dbname_bkup2
```

配置 Oracle RMAN

Oracle RMAN 通道和并行度的配置过程包括指定 Oracle RMAN 备份通道的文件系统目标以及用于备份和恢复操作的总通道数。可通过在所有可用 ZFS Storage Appliance 共享资源上配置 16 个 Oracle RMAN 通道来实现性能优势。配置 Oracle RMAN 通道，将它们平均分配到 Oracle 数据库实例和 RAC 群集中的节点上，并且平均分配到从 ZFS Storage Appliance 中导出的共享资源上。

以下代码片段显示了针对备份集和映像副本执行备份和恢复操作以及将增量合并应用到映像副本的 Oracle RMAN 运行块样例。此样例代码基于以下数据库配置：

- 数据库名称：dbname
- SYSDBA 登录名：sys/welcome
- 扫描地址：ad01-scan
- 备份的服务名称：dbname_bkup

可在单池配置中配置 ZFS Storage Appliance，此设备将在此池中导出用作八个挂载点的八个共享资源。

以下各节中的示例显示了使用备份集和映像副本进行备份和恢复操作的 Oracle RMAN 运行块。在这些示例中，将把四个共享资源的配置的挂载点作为 /zfssa/dbname/backup1 到 /zfssa/dbname/backup4 进行访问。而且，在这些示例针对的配置中，ZFS Storage Appliance 导出四个共享资源，这些共享资源用作 16 个 Oracle RMAN 通道的四个挂载点。

备份集级别 0 备份：

```
run
{<br/>
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';<br/>
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';<br/>
  allocate channel ch01 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';<br/>
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';<br/>
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch09 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';<br/>
  allocate channel ch13 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch14 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch15 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch16 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';<br/>
  configure snapshot controlfile name to<br/>
```

```

'/zfssa/dbname/backup1/snapcf_dbname.f';<br/>
backup as backupset incremental level 0 section size 32g database<br/>
tag 'FULLBACKUPSET_L0' plus archivelog tag 'FULLBACKUPSET_L0';
}

```

备份集级别 1 备份：

```

run
{<br/>
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';<br/>
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';<br/>
  allocate channel ch01 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';<br/>
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch07 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';<br/>
  allocate channel ch08 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch09 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch10 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch11 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch12 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';<br/>
  allocate channel ch13 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';<br/>
  allocate channel ch14 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';<br/>
  allocate channel ch15 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';<br/>
  allocate channel ch16 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';<br/>
  configure snapshot controlfile name to<br/>
'/zfssa/dbname/backup1/snapcf_dbname.f';<br/>
  backup as backupset incremental level 1 database tag<br/>
  'FULLBACKUPSET_L1' plus archivelog tag 'FULLBACKUPSET_L1';
}

```

映像副本备份：

```

run
{<br/>
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';<br/>
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';<br/>
  allocate channel ch01 device type disk connect 'sys/welcome@ad01-<br/>

```

```

scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';<br/>
allocate channel ch02 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';<br/>
allocate channel ch03 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';<br/>
allocate channel ch04 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';<br/>
allocate channel ch05 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup1/%U';<br/>
allocate channel ch06 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup2/%U';<br/>
allocate channel ch07 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup3/%U';<br/>
allocate channel ch08 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup4/%U';<br/>
allocate channel ch09 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';<br/>
allocate channel ch10 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';<br/>
allocate channel ch11 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';<br/>
allocate channel ch12 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';<br/>
allocate channel ch13 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup2/%U';<br/>
allocate channel ch14 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup1/%U';<br/>
allocate channel ch15 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup1' format '/zfssa/dbname/backup4/%U';<br/>
allocate channel ch16 device type disk connect 'sys/welcome@ad01-<br/>
scan/dbname_bkup2' format '/zfssa/dbname/backup3/%U';<br/>
configure snapshot controlfile name to<br/>
'/zfssa/dbname/backup1/snapcf_dbname.f';<br/>
backup incremental level 1 for recover of copy with tag 'IMAGECOPY'<br/>
database;
}

```

增量合并到映像副本：

```

run
{<br/>
  sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';<br/>
  sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';<br/>
  sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';<br/>
  sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';<br/>
  allocate channel ch01 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1';<br/>
  allocate channel ch02 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2';<br/>
  allocate channel ch03 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1';<br/>
  allocate channel ch04 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2';<br/>
  allocate channel ch05 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup1';<br/>
  allocate channel ch06 device type disk connect 'sys/welcome@ad01-<br/>
  scan/dbname_bkup2';<br/>
}

```

```

allocate channel ch07 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch08 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch09 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch10 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch11 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch12 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch13 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch14 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch15 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch16 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
configure snapshot controlfile name to<br/>
'/zfssa/dbname/backup1/snapcf_dbname.f';<br/>
recover copy of database with tag 'IMAGECOPY';
}

```

恢复验证：

```

run
{<br/>
sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';<br/>
sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';<br/>
sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';<br/>
sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';<br/>
allocate channel ch01 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch02 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch03 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch04 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch05 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch06 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch07 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch08 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch09 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch10 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>
allocate channel ch11 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup1';<br/>
allocate channel ch12 device type disk connect 'sys/welcome@ad01-  

scan/dbname_bkup2';<br/>

```

```
allocate channel ch13 device type disk connect 'sys/welcome@ad01-  
<br/>scan/dbname_bkup1';  
<br/>allocate channel ch14 device type disk connect 'sys/welcome@ad01-  
<br/>scan/dbname_bkup2';  
<br/>allocate channel ch15 device type disk connect 'sys/welcome@ad01-  
<br/>scan/dbname_bkup1';  
<br/>allocate channel ch16 device type disk connect 'sys/welcome@ad01-  
<br/>scan/dbname_bkup2';  
<br/>configure snapshot controlfile name to  
<br/>' /zfssa/dbname/backup1/snapcf_dbname.f';  
<br/>restore validate database;  
<br/>}
```

后续步骤

[“Oracle SPARC SuperCluster 备份” \[406\]](#)

为 ZFS Storage Appliance 备份配置 Oracle SPARC SuperCluster

本节包含一些样例脚本，用于说明如何将 ZFS Storage Appliance 连接到 Oracle SPARC SuperCluster。这些脚本用于在单池和双池 ZFS Storage Appliance 配置中支持名为 dbname 的数据库。

配置 SSC：为 ZFS Storage Appliance 备份配置 Oracle SPARC SuperCluster

大致实施步骤

实施步骤如下：

1. 设置在主机上挂载共享资源的目录结构（挂载点）。
2. 更新 /etc/vfstab 以将从 ZFS Storage Appliance 导出的共享资源挂载到相应的挂载点。
3. 如果在重新引导时启用了用来挂载 NFS 共享资源的 NFS 客户机服务，则将自动执行挂载和卸载共享资源的过程。
4. 更新 oranfstab 文件以访问 ZFS Storage Appliance 导出的共享资源。
5. 在主机上挂载共享资源。
6. 更改挂载共享资源的权限以与 ORACLE_HOME 的权限设置匹配。

7. 重新启动 Oracle 数据库实例，以使对 `oranzfstab` 文件进行的更改生效。

具体实施步骤

本节主题：

- “设置在主机上挂载共享资源的目录结构” [425]
- “更新 `/etc/vfstab` 文件” [425]
- “启用 NFS 客户机服务” [426]
- “更新 `oranzfstab` 以访问 ZFS Storage Appliance 导出” [426]
- “在主机上挂载共享资源” [427]
- “设置挂载共享资源的所有权” [427]

设置在主机上挂载共享资源的目录结构

按如下所示设置共享资源在主机上的挂载点：

```
mkdir -p /zfssa/dbname/backup1
mkdir -p /zfssa/dbname/backup2
mkdir -p /zfssa/dbname/backup3
mkdir -p /zfssa/dbname/backup4
```

更新 `/etc/vfstab` 文件

要更新 `/etc/vfstab` 文件，请使用以下选项之一。

注意：下面列表中的 UNIX 换行转义符 (`\`) 表明单行代码在此处换到第二行。将有换行符的行输入 `fstab` 时，请删除 `\` 字符，然后两个行段合并到一行（用空格分隔）。

对于单池配置：

```
192.168.36.200:/export/dbname/backup1 - /zfssa/dbname/backup1 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
192.168.36.200:/export/dbname/backup2 - /zfssa/dbname/backup2 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
192.168.36.200:/export/dbname/backup3 - /zfssa/dbname/backup3 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
192.168.36.200:/export/dbname/backup4 - /zfssa/dbname/backup4 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
```

```
tcp,vers=3,forcedirectio
```

对于双池配置：

```
192.168.36.200:/export/dbname/backup1 - /zfssa/dbname/backup1 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
192.168.36.201:/export/dbname/backup2 - /zfssa/dbname/backup2 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
192.168.36.200:/export/dbname/backup3 - /zfssa/dbname/backup3 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
192.168.36.201:/export/dbname/backup4 - /zfssa/dbname/backup4 \<br/>
nfs - yes rw,bg,hard,nointr,rsiz=1048576,wsiz=1048576,proto= \<br/>
tcp,vers=3,forcedirectio
```

启用 NFS 客户机服务

使用以下命令在 Solaris 11 主机上启用 NFS 客户机服务：

```
svcadm enable -r nfs/client
```

更新 oranfstab 以访问 ZFS Storage Appliance 导出

要更新 oranfstab 文件以访问 ZFS Storage Appliance 导出，请使用以下相应选项。

对于单池配置：

```
server: 192.168.36.200
path: 192.168.36.200
path: 192.168.36.201
path: 192.168.36.202
path: 192.168.36.203
export: /export/dbname/backup1 mount: /zfssa/dbname/backup1
export: /export/dbname/backup2 mount: /zfssa/dbname/backup2
export: /export/dbname/backup3 mount: /zfssa/dbname/backup3
export: /export/dbname/backup4 mount: /zfssa/dbname/backup4
```

对于双池配置：

```
server: 192.168.36.200
path: 192.168.36.200
path: 192.168.36.202
export: /export/dbname/backup1 mount: /zfssa/dbname-2pool/backup1
export: /export/dbname/backup3 mount: /zfssa/dbname-2pool/backup3
server: 192.168.36.201
path: 192.168.36.201
```

```
path: 192.168.36.203
export: /export/dbname/backup2 mount: /zfssa/dbname-2pool/backup2
export: /export/dbname/backup4 mount: /zfssa/dbname-2pool/backup4
```

在主机上挂载共享资源

使用标准的 Solaris mount 命令手动挂载共享资源：

```
# mount /zfssa/dbname/backup1
# mount /zfssa/dbname/backup2
# mount /zfssa/dbname/backup3
# mount /zfssa/dbname/backup4
```

设置挂载共享资源的所有权

更改挂载共享资源的权限设置以与 ORACLE_HOME 的权限设置匹配。在此示例中，用户和组的所有权均设置为 oracle:dba。

1. 输入：
chown oracle:dba /zfssa/dbname/*
2. 使用以下选项之一来重新启动 Oracle 数据库实例，从而获取对 oranfstab 文件所做的更改：

- 一次重新启动一个实例（滚动升级），例如：
- \$ srvctl stop instance -d dbname -i dbname1
- \$ srvctl start instance -d dbname -i dbname1
- \$ srvctl stop instance -d dbname -i dbname2
- \$ srvctl start instance -d dbname -i dbname2
- \$ srvctl stop instance -d dbname -i dbname3
- \$ srvctl start instance -d dbname -i dbname3
- \$ srvctl stop instance -d dbname -i dbname4
- \$ srvctl start instance -d dbname -i dbname4
- \$ srvctl stop instance -d dbname -i dbname5
- \$ srvctl start instance -d dbname -i dbname5
- \$ srvctl stop instance -d dbname -i dbname6
- \$ srvctl start instance -d dbname -i dbname6
- \$ srvctl stop instance -d dbname -i dbname7
- \$ srvctl start instance -d dbname -i dbname7
- \$ srvctl stop instance -d dbname -i dbname8
- \$ srvctl start instance -d dbname -i dbname8

- 重新启动整个数据库，例如：
- `$ srvctl stop database -d dbname`
- `$ srvctl start database -d dbname`

Oracle Intelligent Storage Protocol

Oracle 数据库有一个包含 Oracle Disk Manager (ODM) 的分层体系结构。ODM 提供了一个文件管理模块，允许 Oracle 数据库使用一个本地文件系统、一个原始磁盘分区或 NFS 服务器来存储数据库信息。

为了提高数据库性能，ODM 接口允许 Oracle 数据库在每个 I/O 请求中传递信息。此信息定义了多个与 I/O 关联的属性，例如与 I/O 请求关联的文件类型。这支持以不同方式处理数据文件和数据库日志文件写入操作。

新的 OISP 允许 Oracle Database NFSv4 客户机将 ODM 优化信息传递给 ZFS Storage Appliance 的 NFSv4 服务器。ZFS Storage Appliance 利用 ODM 优化信息简化数据库配置，进一步提高数据库性能。

Oracle Intelligent Storage Protocol 有两个功能：

- 自动设置新数据库文件的最佳文件记录大小
- 自动为每个写入请求使用最佳的写入偏向（ZFS Latency（延迟）或 Throughput（吞吐量））。

设置最佳的文件记录大小

Oracle dNFS 客户机将每个 NFSv4 写入请求的最佳记录大小传递给 ZFS Storage Appliance。ZFS Storage Appliance NFSv4 服务器通过 I/O 请求将记录大小传递给 ZFS 文件系统。ZFS 文件系统然后绕过默认的文件系统记录大小，使用 I/O 请求中传递的记录大小值。只能为新创建的文件设置记录大小。如果一个文件已存在，记录大小将不会改变。

为每个请求使用 ZFS Latency（延迟）或 Throughput（吞吐量）写入模式

Oracle dNFS 客户机将每个 NFSv4 写入请求的最佳写入偏向传递给 ZFS Storage Appliance。ZFS Storage Appliance NFSv4 服务器通过 I/O 请求将写入偏向传递给 ZFS 文件系统。ZFS 文件系统然后绕过默认的文件系统写入偏向，尝试使用 I/O 请求中传递的写入偏向值。根据 ZFS 文件系统的状态，I/O 请求中发送的写入偏向可被忽略。

适用于 Oracle Solaris Cluster 的 Sun ZFS Storage Appliance 网络文件系统插件

Oracle Solaris Cluster (OSC) 是一款适用于 Solaris 操作系统的高可用性群集软件产品。

通过适用于 Oracle Solaris Cluster 的 Sun ZFS Storage Appliance 网络文件系统插件，使包含 Sun ZFS Storage Appliance 的 OSC 能够使用 NFS 协议。Oracle 技术网上适用于 Oracle Solaris Cluster 的 Sun ZFS Storage Appliance 网络文件系统插件中提供了该插件和自述文件。

用于 Oracle Solaris Cluster Geographic Edition 的 Sun ZFS Storage Appliance 插件

Oracle Solaris Cluster Geographic Edition 软件是 Oracle Solaris Cluster 软件的一个分层扩展。Geographic Edition 软件保护应用程序远离意外的中断，使用多个相隔很远的群集，使用一个冗余基础结构来在这些群集站点之间复制数据。此插件使用 Sun ZFS Storage Appliance 远程复制服务协调远程 Oracle Solaris Cluster 站点之间的数据复制。

可通过 Oracle 技术网上的 "Sun NAS Storage" 信息页面获取此插件数据包。

适用于 Oracle Enterprise Manager Grid Control 的 Sun ZFS Storage 管理插件

适用于 Oracle Enterprise Manager Grid Control 的 Sun ZFS Storage 插件对 Sun ZFS Storage Appliance 系列的网格控制器环境提供了一流监视，它具备以下功能：

- 监视 Sun ZFS Storage Appliance
- 收集存储系统信息、配置信息以及可访问的存储组件的性能信息
- 根据阈值以及工具所收集的监视信息发出警报和违规消息
- 提供用于补充 Analytics 的现成可用报告
- 支持通过远程代理进行监视

一旦将某设备配置为通过 Oracle Grid Control 进行监视，就会创建 Analytics 工作表和数据集，以使 Oracle Grid Control 管理员看到通过设备内可用的实时 Analytics 信息所提供的更深层详细信息。

可通过 Oracle 技术网链接获取管理插件。

它与安装指南一起打包提供，网格控制器的管理员和要监视的设备的存储管理员都应阅读该指南。

每个设备中都包含以下两种 [“workflows” \[377\]](#)，分别用于准备系统以进行监视，或者用于删除为监视环境所创建的工件：

- Configure for Oracle Enterprise Manager Monitoring（配置 Oracle Enterprise Manager 监视）
- Unconfigure Oracle Enterprise Manager Monitoring（取消配置 Oracle Enterprise Manager 监视）

可从浏览器用户界面中的[“Maintenance”（维护） > “Workflows”（工作流）](#) [\[377\]](#)页面访问这些工作流。

适用于 Oracle Enterprise Manager Grid Controller 的 Oracle Grid Controller Sun ZFS Storage 管理插件

Configure for Oracle Enterprise Manager Monitoring（配置 Oracle Enterprise Manager 监视）

此工作流用于准备环境以进行监视，或者将该工作流所创建的任何工件重置回其原始状态（如果存储管理员在操作过程中对这些工件进行了更改）。执行此工作流可对系统进行以下更改：

- 将创建对系统具有有限访问权限的 `oracle_agent` [“角色属性” \[129\]](#)，以允许 Oracle Enterprise Manager Grid Controller 代理获取监视所需的信息，而不会对系统进行更改。将创建 `oracle_agent` [第 7 章 用户配置](#)并分配此角色。为保存有关代理何时以及如何访问设备的清晰审计记录，必须使用此角色和用户。
- 将启用 [“Advanced Analytics”（高级 Analytics）](#)，从而为 Sun ZFS Storage Appliance 的所有用户提供扩展的统计信息集。
- 将创建 *Oracle Enterprise Manager* 工作表，以便网格控制器管理员和存储管理员进行沟通。网格控制器监视的所有度量均可在此工作表中查看。

Unconfigure Oracle Enterprise Manager Monitoring（取消配置 Oracle Enterprise Manager 监视）

此工作流将删除 *Configure for Oracle Enterprise Manager Monitoring*（配置 Oracle Enterprise Manager 监视）所创建的工件。具体来说，它：

- 删除 `oracle_agent` 角色和用户，并
- 删除 *Oracle Enterprise Manager* 工作表。

该工作流程不会禁用 "Advanced Analytics" (高级 Analytics) 或为收集数据而激活的任何数据集。

适用于 Sun ZFS Storage Appliance 的 Oracle 虚拟机存储连接插件

Oracle VM 3.0 引入了众多新增功能，Oracle VM Storage Connect 框架就是其中一项。该框架支持 Oracle VM Manager 3.0 直接访问存储服务器和置备资源。通过该框架，您可以注册存储服务器、搜索现有存储资源、创建物理磁盘并将其提供给服务器池以及共享存储系统信息库和虚拟机。

适用于 Sun ZFS Storage Appliance 的 Oracle 虚拟机存储连接插件是 Oracle VM 软件套件的一个组件，该组件支持 Oracle VM 置备和管理 Sun ZFS Storage Appliance 实现虚拟化。该插件可安装在一个或多个 Oracle VM Server 上，并可通过工作流与安装在 ZFSSA 上的一个或多个存储服务器进行通信。

"Oracle Technology Network" (Oracle 技术网) 上提供了此插件和自述文件。

适用于卷影副本服务软件的 Sun ZFS Storage Appliance 提供程序

适用于 Microsoft 操作系统的卷影副本服务 (Volume Shadow Copy Service, VSS) 提供了一个框架，允许在系统上的应用程序仍在写入卷时执行卷备份。VSS 提供了一个一致性接口，允许在更新磁盘上数据的用户应用程序 (VSS 写入程序) 与备份应用程序的用户应用程序 (VSS 请求程序) 之间进行协调。具体来说，VSS 提供了：

- 用于协调应用程序与文件系统活动的备份基础结构
- 用于创建某个时间点合并后副本 (称为 *shadow copies*) 的位置

适用于卷影副本服务软件的 Sun ZFS Storage Appliance 提供程序是一款 VSS 硬件提供程序，它允许 Sun ZFS Storage Appliance 为使用块目标的 Windows 主机创建一致的快照。VSS 会对快照进行协调以确保块数据是一致的。此提供程序与设备上的一组工作流程进行通信，以协调创建在应用程序中看到的快照。它可以通过 iSCSI 和光纤通道运行。

适用于卷影副本服务软件的 Sun ZFS Storage Appliance 提供程序安装在需要此功能和应用程序间协调的主机上。有关此应用程序集成的完整文档以自述文件的形式随下载的组件提供。Oracle 技术网上的 Sun ZFS Storage 7000 软件提供程序和插件补丁程序提供了该提供程序软件和自述文件。有关 VSS 的更多信息可从 Microsoft Web 站点获取，其中包括此<http://msdn.microsoft.com/en-us/library/aa384649> (<http://msdn.microsoft.com/en-us/library/aa384649>)%28VS.85%29.aspx 概述]。

对 Symantec DMP/Storage Foundation 的 FC 支持

- SF – Symantec Storage Foundation 5.1
- SF HA – Storage Foundation High Availability 5.1
- SFCFS/SF Oracle RAC – Storage Foundation Cluster File System/Storage Foundation for Oracle RAC 5.1
- SFCFS/SFCFS Oracle RAC – Storage Foundation Cluster File System/Storage Foundation Cluster File System for Oracle RAC 5.1

针对以下 OS 版本提供对 Symantec Storage Foundation 5.1RP2 及更高版本的 FC 支持

- Solaris 10 SPARC
- Solaris 10 x86
- Linux RedHat5
- Oracle Enterprise Linux (OEL)

请参阅以下位置的 Symantec HCL : <http://www.symantec.com/business/support/index> (<http://www.symantec.com/business/support/index>)?page=content&id=TECH74012

注意以下限制：

- Symantec 要求“必须”安装 7000 ASL，可从以下位置下载：<https://vos.symantec.com/asl>
- Symantec 还要求安装 SF 5.1 VM（修补程序级别为 5.1RP2 或更高版本），可从以下位置下载：<https://vos.symantec.com/patch/matrix>
- Symantec 还要求进行以下 DMP 参数设置（仅适用于“群集化的”7000 系列）：
- `dmp_health_time=0`
- `dmp_path_age=0`
- `dmp_lun_retry_timeout=200`

请参阅 Symantec 的硬件技术说明，其中涉及“群集化的”7000 设置：<http://www.symantec.com/business/support/index> (<http://www.symantec.com/business/support/index>)?page=content&id=TECH47728

Symantec Storage Foundation 5.1SP2 for Windows 支持通过 FC 连接至适用于以下 Windows 版本的 7000 系列：

- Windows Server 2003
- Windows Server 2008
- Windows Server 2008 R2

请参阅以下位置的 SF 5.1SP2 HCL : <http://www.symantec.com/business/support/index>
(<http://www.symantec.com/business/support/index>)?page=content&id=TECH138719

适用于 VMware Site Recovery Manager 的 Sun ZFS Storage 7000 Storage Replication Adapter

适用于 VMware vCenter Site Recovery Manager (SRM) 的 Sun ZFS Storage 7000 Storage Replication Adapter (SRA) 可将 Sun ZFS Storage 7000 设备集成到跨多个站点且需要在受保护的站点服务发生中断时快速恢复的 VMware 部署中。SRA 可插入现有 VMware vCenter SRM 环境，允许在测试和运行恢复计划时通过 VMware vCenter SRM 搜索、测试和故障转移序列管理 Sun ZFS Storage 7000 设备。仅在 VMware vCenter SRM 应用程序内才会使用 SRA。

VMware 管理员需要与负责维护 VMware 数据存储所在设备的 Sun ZFS Storage 7000 设备管理员紧密配合。有关详细信息，请参见 SRA 中包含的《Sun ZFS Storage 7000 SRA for VMware SRM Administration Guide》。

注意：可从 Oracle 技术网下载 SRA。要获取 SRA，需要具备针对 Sun ZFS Storage 7000 设备的有效 Oracle 支持合同。

索引

A

Active

Directory, 228, 228, 228, 229, 229, 229

加入域, 231

加入工作组, 231

B

标头, 26, 26

病毒扫描

配置共享资源的病毒扫描, 223

C

池, 91

存储

将高速缓存设备添加到现有池中, 97, 98

配置存储池, 97

D

Dashboard (显示

板), 46, 46, 48, 49, 50, 53, 54

持续运行 "Dashboard" (显示板), 54

DNS, 239, 239, 239, 241, 241, 241

DTrace, 414

F

FC, 105, 105, 105, 106, 106, 107, 108, 108, 108, 109, 111

FTP

允许对共享资源进行 FTP 访问, 207

G

共享资源, 273, 274, 277

H

HTTP, 208, 208, 209, 209

允许对共享资源进行 HTTP 访问, 209

I

iSCSI

创建 Analytics 工作表, 118

J

警报, 141, 141, 144, 145

添加警报操作, 145, 146

添加阈值警报, 144, 145

K

快照, 303, 304, 304, 304, 307, 308

L

LDAP, 225, 225, 226, 227

添加设备管理员, 227

N

NFS, 108, 109, 111

通过 NFS 共享文件系统, 189

NIC, 69, 75, 75, 87, 88

NIS, 223, 224, 224

从 NIS 添加设备管理员, 224

NTP

BUI 时钟同步, 245

S

设置

- 更改显示的活动统计信息，57
- 更改活动阈值，57

身份映射

- 查看或刷新映射，239
- 配置身份映射，238

SFTP

- 为远程访问配置 SFTP 服务，218
- 允许对共享资源进行 SFTP 访问，217

SMB

- Active Directory 配置，204
- SMB 数据服务配置，205
- 初始配置，203
- 项目和共享资源配置，205

SNMP

- 配置 SNMP 以发送陷阱，253
- 配置 SNMP 以提供设备状态，253

SRP

- iSER 目标配置，118
- SRP 目标配置，124

SSH

- 禁用 root SSH 访问权限，260

- 扩展 LACP 聚合，86

- 添加静态路由，89，89

- 通过拖放创建单个端口接口，84

X

- 显示板，46

- 项目，312

Y

用户

- 为角色添加授权，131，134

- 从角色中删除授权，131，135

- 添加只能查看显示板的用户，131

- 添加管理员，130，133

- 添加角色，131，134

- 远程复制，327，330，330，351

T

TFTP

- 允许对共享资源进行 TFTP 访问，220

W

网络

- 为群集控制器创建不带 VLAN ID 的 VNIC，87

- 为群集控制器创建带有相同 VLAN ID 的 VNIC，88

- 修改接口，84

- 创建 InfiniBand 分区数据链路和接口，86

- 创建 LACP 聚合链路接口，84

- 创建仅使用链路状态故障检测的 IPMP 组，85

- 创建使用基于探测和链路状态故障检测的 IPMP 组，85

- 创建单个端口接口，83

- 删除静态路由，89，89

- 将多宿主属性更改为 strict，90

- 扩展 IPMP 组，86