

Oracle® Big Data Appliance

Configuration Worksheets

Release 1 (1.0.3)

E26163-05

June 2012

The *Oracle Big Data Appliance Configuration Worksheets* identify the information and decisions required when installing Oracle Big Data Appliance. This document contains the following topics:

- [Configuration Process Overview](#)
- [Network Configuration Worksheets](#)
- [Software Configuration Worksheets](#)
- [Sample Installation Template](#)
- [Documentation Accessibility](#)

Configuration Process Overview

The configuration worksheets will help you plan your installation of Oracle Big Data Appliance with the network and database administrators and your Oracle representative. It is important that you complete the worksheets and provide them to your Oracle representative before receiving your Oracle Big Data Appliance to avoid delaying the installation.

Your Oracle representative uses the information entered in these worksheets to create the Oracle Big Data Appliance Installation Template. Use the Installation Template to confirm that the configuration settings of your Oracle Big Data Appliance were entered correctly. You can make site-specific adjustments to the Installation Template in consultation with your Oracle representative. Many of these settings cannot be changed after the appliance is operational, so make your decisions carefully.

Note:

- Oracle Big Data Appliance uses Cloudera's Distribution including Apache Hadoop (CDH). A Hadoop cluster on Oracle Big Data Appliance is called a CDH cluster.
- The terms **appliance** and **rack** refer to Oracle Big Data Appliance.

To configure Oracle Big Data Appliance:

1. Read this document and the *Oracle Big Data Appliance Owner's Guide* to understand the networking requirements.
2. Identify your current management network.
3. Identify your current client access network.

4. If you have multiple appliances, then choose between connecting them into a single CDH cluster or creating an independent CDH cluster on each appliance.
5. Complete the PDF version of the configuration worksheets. If you are configuring multiple appliances as a single CDH cluster, then complete just one copy of the worksheets. Otherwise, complete a copy of the worksheets for every set of appliances that will be configured as a single CDH cluster.
6. Give the completed configuration worksheets to your Oracle representative. The information in these worksheets is used to generate the Installation Template, which your Oracle representative will give to you.
7. Verify that the information in the Installation Template is correct. To make any adjustments, contact your Oracle representative.
8. Configure your existing network to use the new IP addresses and host names supplied in the Installation Template.
9. Review the network and IP address requirements described in the *Oracle Big Data Appliance Owner's Guide*.
10. Run the network connections to the planned location for Oracle Big Data Appliance.
11. Inform your Oracle representative when you have completed these steps.

Note: The network administrator should be available on-site during the installation of Oracle Big Data Appliance (1 to 2 days). Otherwise, a problem with the customer network may cause extended delays.

See Also:

- *Oracle Big Data Appliance Owner's Guide* for more information about the configuration settings.
- *Oracle Big Data Appliance Site Checklists* for site preparations before installation.

Network Configuration Worksheets

The network configuration worksheets identify the naming conventions for the hardware and the IP addresses used on your network:

- [General Properties for the Rack](#)
- [Network Properties for the Rack](#)
- [Network Properties for the Servers](#)
- [Network Properties for the Switches](#)
- [Network Properties for the PDUs](#)

General Properties for the Rack

In [Table 1](#), enter the names for the hardware components and other basic information.

Using Standardized Host Names

All racks in a CDH cluster must have the same CDH cluster name, which is used in the assignment of standardized host names for all Oracle Big Data Appliance servers. By default, the host name for all 18 servers in the rack is in this format:

rack_namenodeNN.domain

Where:

- *rack_name* by default is *cluster_nameM*, and *M* is the index of the rack in the cluster.
- *NN* is the position number of the server in the rack (01-18).
- *domain* is the domain name.

For example, if the cluster name is *bda* and the domain is *example.com*, then the fully qualified host name of the server at the bottom of the first rack is *bda1node01.example.com*. For the top server in the third rack of this cluster, the host name is *bda3node18.example.com*.

The host names must have fewer than 38 characters. You use them to address the servers over the client network interface.

The host names on the other networks have the following formats. You use these formats when connecting over the other networks:

- For short host names over the administration network:
rack_namenodeNN-adm
- For the private InfiniBand network host names:
rack_namenodeNN-priv
- For the Oracle Integrated Lights Out Manager (ILOM) host names:
cluster_nameMnodeNN-c
- For the switch host names:
cluster_nameMsw-ibN
where *N* is 1, 2 or 3, depending on the switch location in the rack.

Using Customized Host Names

If you do not want to use the cluster name in the host names, then you must provide a rack name. You also can change the suffixes used for the different network interfaces.

To assign a unique host name to each server, provide your Oracle representative with a list of all individual host names. You must also enter the domain name and time zone fields in [Table 1](#).

Table 1 General Rack Configuration Worksheet

Property	Example	Setting
Customer Name ¹	Example Inc	
CDH Cluster Name ¹	bda	
Is this the only rack or the primary rack in the cluster? ¹	Yes or No	
Rack Name	bda1	
Server Base Name	node	
Admin Access Suffix	-adm	
Private Name Suffix	-priv	
ILOM Name Suffix	-c	
Switch Base Name	sw	
Domain Name ¹	example.com	
Time Zone ¹	America/New_York	

¹ Required. All other properties in this worksheet have default values.

Network Properties for the Rack

In [Table 2](#), enter up to four IP addresses for the DNS and NTP servers and up to four search domains.

Table 2 Network Configuration Worksheet for the Rack

Property	Setting
DNS Servers (1-4)	
NTP Servers (1-4)	
Search Domains (1-4)	

Network Properties for the Servers

In [Table 3](#), enter the network properties for the individual servers.

Each server in the appliance is assigned an IP address for the administrative 1 gigabit Ethernet (1 GbE) network, the private InfiniBand network, the 10 GbE client access network, and Oracle ILOM. The IP addresses of the 18 servers are assigned sequentially beginning with the Start IP for node01.

If you cannot allocate 18 sequential IP addresses for a network, then provide your Oracle representative with a list of 18 IP addresses instead of just the starting IP address.

The administrative and Oracle ILOM IP addresses are on the same subnet and use the same netmask and gateway.

Table 3 Network Configuration Worksheet for the Servers

Property	Setting
Administrative Network: First of 18 sequential IP addresses	
Administrative Network Netmask	
Administrative Network Gateway	
Client Network: First of 18 sequential IP addresses	
Client Network Netmask	
Client Network Gateway	
ILOM Network: First of 18 sequential IP addresses	
Are you connecting this rack to another rack through the InfiniBand fabric, such as another Oracle Big Data Appliance rack or an Oracle Exadata Database Machine?	
Private InfiniBand Network: First of 18 sequential IP addresses Optional: Defaults to 192.168.10.1 ¹	
Private InfiniBand Network Netmask Optional: Defaults to 255.255.255.0 ¹	

¹ If you are connecting this rack to another rack through the InfiniBand fabric, see the following note.

Note: If you are connecting this rack to another rack through the InfiniBand fabric, then ensure that:

- The InfiniBand IP addresses of all servers are unique, including the servers in other Oracle engineered systems.
- All InfiniBand IP addresses are on the same network.

For example, if you connect an Oracle Big Data Appliance rack to Oracle Exadata Database Machine, then you must use the same netmask on the InfiniBand networks for both systems. Moreover, after you apply this netmask to the InfiniBand IP addresses of the Exadata compute nodes, the Exadata storage servers, and the Oracle Big Data Appliance servers, all IP addresses are in the same subnet.

Network Properties for the Switches

In [Table 4](#), enter the IP addresses of the keyboard-video-mouse (KVM) Switch, the Cisco Catalyst Ethernet Switch, and the first of three Sun InfiniBand switches.

Table 4 Network Configuration Worksheet for the Switches

Property	Setting
KVM Switch IP	
Cisco Switch IP	
InfiniBand Switch IP: First of three sequential IP addresses	

Network Properties for the PDUs

In [Table 5](#), enter the IP addresses for the two power distribution units (PDUs).

Table 5 Network Configuration Worksheet for the PDUs

Property	Setting
PDU A	
PDU B	

Software Configuration Worksheets

There are several optional software components for Oracle Big Data Appliance, which you can choose to configure and activate.

- [Installed Components](#)
- [Auto Service Request](#)
- [Users and Groups](#)
- [Cloudera Manager Email Alerts](#)

Note: If you are extending a CDH cluster and these worksheets are not for the primary rack, then complete only [Table 8](#). Leave the other Software Configuration Worksheets blank.

See Also: *Oracle Big Data Appliance Owner's Guide* for a description of the software configuration options.

Installed Components

In [Table 6](#), identify the applications to activate and configure.

Optional Software

Oracle Big Data Connectors and Oracle NoSQL Database are optional components. You must activate them now if you plan to use them. Otherwise, you can save disk space for other uses by not activating them. Disk space allocated to Oracle NoSQL Database is not available for the Hadoop Distributed File System (HDFS).

Oracle NoSQL Database Community Edition is included in the license for Oracle Big Data Appliance.

Oracle Big Data Connectors requires a separate license. You must have this license to install the connectors on Oracle Big Data Appliance. If you have a license, you can decide whether to configure Oracle Data Integrator Agent during the initial software installation, so that it is up and running immediately.

External NameNode Backup

Oracle Big Data Appliance is configured to have four copies of the critical NameNode data spread across four disks on two servers. This default of four copies inside Oracle Big Data Appliance and no external copy is the Oracle recommended configuration. It is sufficient for most disaster recovery scenarios. Oracle recommends that you keep this default backup configuration.

Nonetheless, if you want to maintain a copy on a separate server outside of Oracle Big Data Appliance, you can provide an external network file system (NFS) mount point for that purpose. Oracle Big Data Appliance then maintains two copies of the NameNode data on the appliance and one copy on the external mount point. Be aware that when you provide an external mount point, the number of copies inside Oracle Big Data Appliance is reduced from four copies to two copies.

Note: To maintain an external copy of the NameNode data, you must manually define a mount point *before* you install the Oracle Big Data Appliance software.

To set up an external NFS mount point for backing up the NameNode data:

1. Log in to the NFS host. You must have `root` privileges.
2. Create the `hdfs` user in the `hadoop` group. The `hdfs` UID must be the same as the `hdfs` UID on Oracle Big Data Appliance, and the `hadoop` GID must be the same as the `hadoop` GID on Oracle Big Data Appliance.
3. Create the top-level directory for the backup files.
4. As the `root` user, create a subdirectory with the same name as the cluster under the new directory path.
5. Verify that this subdirectory is owned by `root`.
6. As the `hdfs` user, create two subdirectories named `nn` and `snn` under the new subdirectory.
7. Verify that these subdirectories are owned by user `hdfs` and group `hadoop`.
8. Log in as the `root` user to `node01` of Oracle Big Data Appliance.
9. Create a mount point by issuing a command like the following:

```
mount nfs-host:/backup/namenode /mnt/nfs-backup
```

In this example, `NFS-HOST` is the name of the external NFS host, `/backup/namenode` is the top-level directory for the backup files, and `/mnt/nfs-backup` is the mount point on Oracle Big Data Appliance.

10. In [Table 6](#), enter the full path of the top-level directory from Step 3.

For example, if you enter the NFS directory as nfs-host:/backup/namenode and the cluster name is bda, then:

- The /backup/namenode/bda directory must exist on NFS-HOST and be owned by root.
- The /backup/namenode/bda/nn and /backup/namenode/bda/snn directories must exist on NFS-HOST and be owned by hdfs in group hadoop.

The directory structure on NFS-HOST looks like this:

```
/backup
  /namenode
    /bda
      /nn
      /snn
```

Note: Complete this worksheet for the only rack or the primary rack in a CDH cluster.

Table 6 Software Components Configuration Worksheet

Property	Example	Setting
Are Oracle Big Data Connectors licensed?	Yes or No	
Configure Oracle Data Integrator Agent?	Yes or No	
Install Oracle NoSQL Database Community Edition?	Yes or No	
Total disk space to allocate for Oracle NoSQL Database (in terabytes)	0, 54, or 108	
External backup Network File System (NFS) directory	nfs-host:/backup/namenode	

Auto Service Request

In [Table 7](#), enter the configuration settings for Auto Service Request (ASR).

ASR monitors the health of Oracle Big Data Appliance hardware and automatically submits a service request when it detects a fault. Although you can opt out of this program, Oracle recommends that you enable ASR.

ASR Manager must be configured to run on a separate server outside of Oracle Big Data Appliance. Software on Oracle Big Data Appliance must be able to connect to ASR Manager and route to the Internet, either directly or through a proxy, to send event information that automatically opens service requests.

Note: Complete this worksheet for the only rack or the primary rack in a CDH cluster.

Table 7 Auto Service Request Configuration Worksheet

Property	Example	Setting
Enable Auto Service Request?	Yes or No	
ASR Manager host address	asr-host.example.com	
ASR Manager port number	162	
ASR Server root password	--	

Users and Groups

The installation software defines users, groups, and passwords. If you do not provide the passwords in [Table 8](#) for the configuration files, then you must enter them manually during the installation.

The Oracle IDs must match those of a connected Oracle Exadata Database Machine to support Network File System (NFS) protocol between the two systems.

Note: If you are extending a CDH cluster and these worksheets are not for the primary rack, then complete only the password fields in [Table 8](#). Leave the ID fields blank.

Table 8 Users and Groups Configuration Worksheet

Description	Example	Value
Cloudera Manager admin password ¹	--	
root operating system password ¹	--	
oracle operating system password ¹	--	
oracle user ID	1000	
oinstall group ID	1001	
dba group ID	1002	
MySQL administration password ¹	--	
MySQL password for Oracle Data Integrator ¹	--	

¹ Required for all racks.

Cloudera Manager Email Alerts

In [Table 9](#), provide information about the email server (SMTP) on the network.

This information enables Cloudera Manager to send email alerts when it detects a problem in the CDH cluster. If this information is omitted, then no email alerts will be sent.

Note: Complete this worksheet for the only rack or the primary rack in a CDH cluster.

Table 9 Email Configuration Worksheet

Property	Example	Value
Email Server host address	router.example.com	
Email Server port number	25	
Email Server user name	smtp.user	
Email Server password	--	
Does the email server use SSL?	Yes or No	
Email alert recipients	admin@example.com	

Sample Installation Template

Following is an example of an Installation Template generated using the example values shown in these worksheets.

Network Configuration

General

Property	Value
Cluster Name	bda
Rack Name	bda1
Primary Appliance	Yes
Country	America
Timezone	America/New_York
Domain	example.com
DNS Server	172.16.100.45
NTP Server	172.16.215.62
Search Domains	example.com us.example.com

Server Network Info

Type	Netmask	Gateway
Administrative - eth0	255.255.255.0	10.18.113.1
Private - bondib0	255.255.255.0	--
Client Access - bondeth0	255.255.255.0	10.18.114.1

Server Position	eth0 hostname	eth0 IP address	bondib0 IP address	...
01	bda1node01-adm	10.18.113.10	bda1node01-priv	...
02	bda1node02-adm	10.18.113.11	bda1node02-priv	...
03	bda1node03-adm	10.18.113.12	bda1node03-priv	...

Server Position	eth0 hostname	eth0 IP address	bondib0 IP address	...
...
17	bda1node17-adm	10.18.113.26	bda1node17-priv	...
18	bda1node18-adm	10.18.113.27	bda1node18-priv	...

Switches

Type	Hostname	IP Address
KVM Switch	bda1sw-kvm	10.18.113.143
Cisco Switch	bda1sw-ip	10.18.113.144
NM2-36p Spine Switch	bda1sw-ib1	10.18.113.145
NM2-GW Leaf Switch #1	bda1sw-ib2	10.18.113.146
NM2-GW Leaf Switch #2	bda1sw-ib3	10.18.113.147

PDUs

Type	Hostname	IP Address
PDU A	bda1-pdu1	10.18.113.151
PDU B	bda1-pdu2	10.18.113.152

Software Configuration

Property	Value
Oracle Big Data Connectors	are being licensed
Oracle Data Integrator Agent	will be configured and started
Oracle NoSQL Database	will be installed with 54 TB total diskspace
External NFS Directory	no external NFS mount point requested

Property	Value
Oracle Auto Service Requests	will be configured
Oracle ASR server	asr-host.example.com:162
Oracle ASR Server root password	password

Property	Value
Cloudera Manager admin password	password
root operating system password	password
oracle operating system password	password

Property	Value
MySQL administration password	password
MySQL password for Oracle Data Integrator	password

Property	Value	Property	Value	Property	Value
oracle user ID	1000	oinstall group ID	1001	dba group ID	1002

Property	Value
Email server (SMTP) host	router.example.com:25 does not use SSL
Email server (SMTP) user name	smtp.user
Email server (SMTP)	password
Email alert recipients	admin@example.com

Documentation Accessibility

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

Access to Oracle Support

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

Oracle Big Data Appliance Configuration Worksheets, Release 1 (1.0.3)
E26163-05

Copyright © 2012, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

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.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

Cloudera, Cloudera CDH, and Cloudera Manager are registered and unregistered trademarks of Cloudera, Inc.