

Oracle Exadata Database Machine Extending and Multi-Rack Cabling Guide



21.2
F29251-15
February 2022

ORACLE®

Copyright © 2008, 2022, Oracle and/or its affiliates.

Primary Authors: Peter Fusek, Glenn Maxey

Contributing Authors: Craig Brown, Henry Chow, Holger Leister

Contributors: Doug Archambault, Leo Agranonik, Nilesh Choudhury, Jaime Figueroa, Roger Hansen, Leslie Keller, Frank Kobylanski, René Kundersma, Yang Liu, Juan Loaiza, Barb Lundhild, Philip Newlan, Dan Norris, Michael Nowak, Gavin Parish, Hector Pujol, Darryl Presley, Ashish Ray, Richard Scales, Oliver Sharwood, Jia Shi, Kesavan Srinivasan, Krishnadev Telikicherla, Cliff Thomas, Alex Tsukerman, Kothanda Umamageswaran, Doug Utzig, Zheren Zhang

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, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs) and Oracle computer documentation or other Oracle data delivered to or accessed by U.S. Government end users are "commercial computer software" or "commercial computer software documentation" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, reproduction, duplication, release, display, disclosure, modification, preparation of derivative works, and/or adaptation of i) Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs), ii) Oracle computer documentation and/or iii) other Oracle data, is subject to the rights and limitations specified in the license contained in the applicable contract. The terms governing the U.S. Government's use of Oracle cloud services are defined by the applicable contract for such services. 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, Java, and MySQL are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Inside 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, Epyc, and the AMD 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 about 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 unless otherwise set forth in an applicable agreement between you and Oracle. 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, except as set forth in an applicable agreement between you and Oracle.

Contents

Preface

Audience	vii
Documentation Accessibility	vii
Diversity and Inclusion	vii
Related Documentation	viii
Conventions	ix

1 Preparing to Extend Oracle Exadata Database Machine

1.1	About Extending Exadata Database Machine	1-1
1.2	Reviewing the Safety Precautions	1-5
1.3	Reviewing the Cable Precautions	1-5
1.4	Estimating Cable Path Lengths	1-6
1.5	Bundling Cables	1-6
1.5.1	Floor and Underfloor Delivery of Cables	1-7
1.6	Reviewing the Cable Management Arm Guidelines	1-7
1.7	Obtaining Current Configuration Information	1-8
1.8	Preparing the Network Configuration	1-9
1.9	Moving Audit and Diagnostic Files	1-10
1.10	Reviewing Release and Patch Levels	1-10
1.11	Performing Preliminary Checks	1-11
1.12	Preparing to Add Servers	1-11

2 Extending the Hardware

2.1	Extending an Eighth Rack	2-1
2.1.1	Eighth Rack Expansion Options	2-2
2.1.1.1	X9M-2 Eighth Rack Upgrade Options	2-2
2.1.1.2	X7-2, X8-2, and X8M-2 Eighth Rack Upgrade Options	2-2
2.1.1.3	X6M-2 Eighth Rack Upgrade Options	2-3
2.1.1.4	Eighth Rack X4-2 and X5-2 Expansion Options	2-3
2.1.2	Eighth Rack Expansion Procedures	2-3

2.1.2.1	Reviewing and Validating Current Configuration of Eighth Rack Oracle Exadata Database Machine	2-4
2.1.2.2	Activating Database Server Cores in Oracle Exadata Database Machine Eighth Rack	2-5
2.1.2.3	Activating Storage Server Cores and Disks in Oracle Exadata Database Machine Eighth Rack	2-6
2.1.2.4	Creating Additional Grid Disks in Oracle Exadata Database Machine Eighth Rack	2-7
2.1.2.5	Adding Grid Disks to Oracle ASM Disk Groups in Oracle Exadata Database Machine Eighth Rack	2-13
2.1.2.6	Validating Expansion of Oracle Exadata Database Machine	2-15
2.2	Extending Elastic Configurations	2-17
2.2.1	Removing the Doors	2-18
2.2.2	Adding New RDMA Network Fabric Switches	2-18
2.2.2.1	Adding a RoCE Network Fabric Switch (Cisco Nexus 9336C-FX2)	2-18
2.2.2.2	Adding an InfiniBand Network Fabric Switch (Sun Datacenter InfiniBand Switch 36)	2-19
2.2.3	Adding New Servers	2-21
2.2.3.1	Preparing to Install New Servers	2-22
2.2.3.2	Installing the Rack Assembly	2-22
2.2.3.3	Installing the Server	2-23
2.2.4	Cabling Database Servers	2-24
2.2.5	Cabling Storage Servers	2-29
2.2.6	Closing the Rack	2-31
2.3	Extending a Rack by Adding Another Rack	2-31
2.3.1	Overview of Adding Another Rack to an Existing System	2-32
2.3.2	Cabling Two Racks Together	2-33
2.3.2.1	Cabling Two RoCE Network Fabric Racks Together with No Down Time	2-33
2.3.2.2	Cabling Two RoCE Network Fabric Racks Together with Down Time	2-158
2.3.2.3	Cabling Two InfiniBand Network Fabric Racks Together	2-178
2.3.3	Cabling Several Racks Together	2-179
2.3.3.1	Cabling Several RoCE Network Fabric Racks Together using Oracle Exadata System Software Release 20.1.0 or Later	2-180
2.3.3.2	Cabling Several RoCE Network Fabric Racks Together using Oracle Exadata System Software Release 19.3 or Earlier	2-186
2.3.3.3	Cabling Several InfiniBand Network Fabric Racks Together	2-199

3 Configuring the New Hardware

3.1	Changing the Interface Names	3-1
3.2	Setting Up New Servers	3-2
3.2.1	Configuring Servers Using OEDA	3-3
3.2.2	Configuring New Servers Manually	3-5

3.3	Setting up a New Rack	3-6
3.4	Setting User Equivalence	3-7
3.5	Starting the Cluster	3-9
3.6	Adding Grid Disks to Oracle ASM Disk Groups	3-10
3.7	Adding Servers to a Cluster	3-14
3.8	Configuring Cell Alerts for New Oracle Exadata Storage Servers	3-21
3.9	Adding Oracle Database Software to the New Servers	3-22
3.10	Adding Database Instance to the New Servers	3-24
3.11	Returning the Rack to Service	3-24

4 Multi-Rack Cabling Tables for Oracle Exadata Rack X9M

4.1	Understanding Multi-Rack Cabling for X9M Racks	4-1
4.2	Preparing for Multi-Rack Cabling with X9M Racks	4-4
4.3	Two-Rack Cabling for X9M Racks	4-6
4.4	Three-Rack Cabling for X9M Racks	4-9
4.5	Four-Rack Cabling for X9M Racks	4-13
4.6	Five-Rack Cabling for X9M Racks	4-18
4.7	Six-Rack Cabling for X9M Racks	4-24
4.8	Seven-Rack Cabling for X9M Racks	4-31
4.9	Eight-Rack Cabling for X9M Racks	4-40

5 Multi-Rack Cabling Tables for Oracle Exadata Rack X8M

5.1	Understanding Multi-Rack Cabling for X8M Racks	5-1
5.1.1	Preparing for Multi-Rack Cabling with X8M Racks	5-4
5.1.2	Cabling Multiple Exadata Database Machine X8M Racks	5-6
5.2	Two-Rack Cabling for X8M Racks	5-7
5.3	Three-Rack Cabling for X8M Racks	5-9
5.4	Four-Rack Cabling for X8M Racks	5-13
5.5	Five-Rack Cabling for X8M Racks	5-18
5.6	Six-Rack Cabling for X8M Racks	5-22
5.7	Seven-Rack Cabling for X8M Racks	5-28
5.8	Eight-Rack Cabling for X8M Racks	5-35

6 Multi-Rack Cabling Tables for Oracle Exadata Rack Models with InfiniBand Network Fabric (X2 to X8)

6.1	Understanding Multi-Rack Cabling for Racks with InfiniBand Network Fabric	6-1
6.1.1	Preparing for Multi-Rack Cabling with InfiniBand Network Fabric	6-5
6.1.2	Cabling Oracle Exadata Quarter Racks and Oracle Exadata Eighth Racks with InfiniBand Network Fabric	6-8

6.2	Two-Rack Cabling with InfiniBand Network Fabric	6-10
6.3	Three-Rack Cabling with InfiniBand Network Fabric	6-12
6.4	Four-Rack Cabling with InfiniBand Network Fabric	6-14
6.5	Five-Rack Cabling with InfiniBand Network Fabric	6-17
6.6	Six-Rack Cabling with InfiniBand Network Fabric	6-20
6.7	Seven-Rack Cabling with InfiniBand Network Fabric	6-24
6.8	Eight-Rack Cabling with InfiniBand Network Fabric	6-28

Preface

This guide describes how to extend Oracle Exadata Database Machine and cable multiple racks together. It includes information about the cables, new server installation, and cabling tables.

- [Audience](#)
- [Documentation Accessibility](#)
- [Diversity and Inclusion](#)
- [Related Documentation](#)
- [Conventions](#)

Audience

This guide is intended for Oracle Exadata Database Machine customers and those responsible for data center site planning, configuration, and maintenance of Exadata Database Machine.

Documentation Accessibility

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

Access to Oracle Support

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

Diversity and Inclusion

Oracle is fully committed to diversity and inclusion. Oracle respects and values having a diverse workforce that increases thought leadership and innovation. As part of our initiative to build a more inclusive culture that positively impacts our employees, customers, and partners, we are working to remove insensitive terms from our products and documentation. We are also mindful of the necessity to maintain compatibility with our customers' existing technologies and the need to ensure continuity of service as Oracle's offerings and industry standards evolve. Because of these technical constraints, our effort to remove insensitive terms is ongoing and will take time and external cooperation.

Related Documentation

In addition to the Exadata Database Machine documentation set, the following guides contain hardware information for Exadata Database Machine:

- *Oracle Exadata Database Machine System Overview*
- *Oracle Exadata Database Machine Security Guide*
- *Oracle Engineered System Safety and Compliance Guide, Compliance Model No.: ESY27*
- *Oracle Exadata Database Machine Extending and Multi-Rack Cabling Guide*
- *Oracle Exadata Database Machine Maintenance Guide*
- *Oracle Exadata System Software User's Guide*
- *Oracle Exadata X8-2 Database Server Service Manual* at http://docs.oracle.com/cd/E93359_01/html/E93386/index.html
- *Oracle Exadata Storage Server X8-2 EF,HC, and XT Service Manual* at https://docs.oracle.com/cd/E93361_01/html/E93395/index.html
- *Oracle Exadata Database Server X8-8 Service Manual* at http://docs.oracle.com/cd/E93360_01/html/E96097/index.html
- *Oracle Server X7-2 Service Manual* at http://docs.oracle.com/cd/E72435_01/html/E72445/index.html
- *Oracle Server X7-2L Service Manual* at http://docs.oracle.com/cd/E72463_01/html/E72474/index.html
- *Oracle Server X7-8 Service Manual* at http://docs.oracle.com/cd/E71925_01/html/E71936/index.html
- *Oracle Server X6-2 Service Manual* at http://docs.oracle.com/cd/E62159_01/html/E62171/index.html
- *Oracle Server X6-2L Service Manual* at http://docs.oracle.com/cd/E62172_01/html/E62184/index.html
- *Oracle Server X5-2 Service Manual* at http://docs.oracle.com/cd/E41059_01/html/E48312/napsm.html
- *Oracle Server X5-2L Service Manual* at http://docs.oracle.com/cd/E41033_01/html/E48325/cnpsm.html#scrolltoc
- *Sun Server X4-8 Service Manual* at http://docs.oracle.com/cd/E40591_01/html/E40317/index.html
- *Sun Server X4-2 Service Manual* at http://docs.oracle.com/cd/E36975_01/html/E38045/gentextid-14757.html#scrolltoc
- *Sun Server X4-2L Service Manual* at http://docs.oracle.com/cd/E36974_01/html/E38145/gentextid-14728.html#scrolltoc
- *Sun Server X3-2 (formerly Sun Fire X4170 M3) Service Manual* at http://docs.oracle.com/cd/E22368_01/html/E27242/gentextid-14840.html#scrolltoc
- *Sun Server X3-2L (formerly Sun Fire X4270 M3) Service Manual* at http://docs.oracle.com/cd/E23393_01/html/E27229/gentextid-14804.html#scrolltoc

- *Sun Server X2-8 (formerly Sun Fire X4800 M2) Service Manual* at http://docs.oracle.com/cd/E20815_01/html/E20819/index.html
- *Sun Fire X4800 Server Service Manual* at <http://docs.oracle.com/cd/E19140-01/html/821-0282/index.html>
- *Sun Fire X4270 M2 Server Service Manual* at <http://docs.oracle.com/cd/E19245-01/E21671/index.html>
- *Sun Fire X4170 M2 Server Service Manual* at <http://docs.oracle.com/cd/E19762-01/E22369-02/index.html>
- *Sun Fire X4170, X4270, and X4275 Servers Service Manual* at <http://docs.oracle.com/cd/E19477-01/820-5830-13/index.html>
- *Sun Datacenter InfiniBand Switch 36 Firmware Version 2.1 Documentation* at http://docs.oracle.com/cd/E36265_01/index.html
- *Sun Datacenter InfiniBand Switch 36 Firmware Version 2.2 Documentation* at http://docs.oracle.com/cd/E76424_01/index.html
- *Sun Flash Accelerator F20 PCIe Card User's Guide* at <http://docs.oracle.com/cd/E19682-01/E21358/index.html>
- *Sun Flash Accelerator F40 PCIe Card User's Guide* at http://docs.oracle.com/cd/E29748_01/html/E29741/index.html
- *Sun Flash Accelerator F80 PCIe Card User's Guide* at http://docs.oracle.com/cd/E41278_01/html/E41251/index.html
- *Oracle Flash Accelerator F160 PCIe Card User Guide* at http://docs.oracle.com/cd/E54943_01/html/E54947/index.html
- *Oracle Flash Accelerator F320 PCIe Card User Guide* at http://docs.oracle.com/cd/E65386_01/html/E65387/index.html
- *Oracle Flash Accelerator F640 PCIe Card User Guide* at https://docs.oracle.com/cd/E87231_01/html/E87233/index.html
- *Sun Storage 6 Gb SAS PCIe RAID HBA Documentation* at <http://docs.oracle.com/cd/E19221-01/>
- *Oracle Storage 12 Gb/s SAS PCIe RAID HBA, Internal Documentation Library* at http://docs.oracle.com/cd/E52363_01/index.html
- *Oracle Integrated Lights Out Manager (ILOM) Documentation* at <http://www.oracle.com/goto/ilom/docs>
- "Cisco Catalyst 4948E and 4948E-F Ethernet Switches Data Sheet" at https://www.cisco.com/c/en/us/products/collateral/switches/catalyst-4948e-ethernet-switch/data_sheet_c78-598933.html
- "Cisco Nexus 9300-EX and 9300-FX Platform Switches Data Sheet" at <https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-736651.html>

Conventions

The following text conventions are used in this document:

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

1

Preparing to Extend Oracle Exadata Database Machine

Before extending any rack hardware, review the safety precautions and cabling information, and collect information about the current rack in this section.

- [About Extending Exadata Database Machine](#)
You can extend Exadata Database Machine either by adding servers to the current configuration or by cabling together multiple racks.
- [Reviewing the Safety Precautions](#)
Before upgrading Oracle Exadata Database Machines, read *Important Safety Information for Sun Hardware Systems* included with the rack.
- [Reviewing the Cable Precautions](#)
- [Estimating Cable Path Lengths](#)
- [Bundling Cables](#)
- [Reviewing the Cable Management Arm Guidelines](#)
Review the following cable management arm (CMA) guidelines before routing the cables.
- [Obtaining Current Configuration Information](#)
- [Preparing the Network Configuration](#)
When adding additional servers to your rack, you will need IP address and the current network configuration settings.
- [Moving Audit and Diagnostic Files](#)
- [Reviewing Release and Patch Levels](#)
When adding new servers to a rack, you must match the installed operating system version and software releases.
- [Performing Preliminary Checks](#)
- [Preparing to Add Servers](#)

1.1 About Extending Exadata Database Machine

You can extend Exadata Database Machine either by adding servers to the current configuration or by cabling together multiple racks.

Here are considerations when extending Exadata Database Machine:

- You can extend Exadata Database Machine from a fixed or custom configuration to another configuration by adding any combination of database or storage servers up to the allowed maximum.
- You can cable together multiple Exadata Database Machine racks subject to the following:
 - You can cable together different rack models. For example, you can cable together an X8-2 rack and an X7-2 rack.

- All racks that are cabled together in a multi-rack configuration must use the same RDMA Network Fabric. That is, all racks must use RoCE Network Fabric, or all racks must use InfiniBand Network Fabric.

You cannot have a mixture of racks using RoCE Network Fabric and InfiniBand Network Fabric. For example, you cannot cable together an X8-2 rack and an X9M-2 rack.

- All racks that are cabled together in a multi-rack configuration have the same database server hardware architecture. That is, all racks must use 2-socket database servers, or all racks must use 8-socket database servers.

You cannot have a mixture of racks using 2-socket and 8-socket database servers. For example, you cannot cable together an X9M-2 rack and an X9M-8 rack.

- Prior to extending a system across multiple racks, you must acquire the appropriate RDMA Network Fabric switches and transceivers.
- When extending Exadata Database Machine Eighth Rack with Oracle Exadata Storage Expansion Rack, Oracle recommends using separate disk groups for the disks in each rack.

Multiple Exadata Database Machines can run as separate environments while sharing the RDMA Network Fabric. If you are planning to utilize multiple Exadata Database Machines in this manner, then note the following:

- All servers on the RDMA Network Fabric must have a unique IP address. When Exadata Database Machine is deployed, the default network is 192.168.10.1. You must modify the IP addresses before re-configuring the RDMA Network Fabric. Failure to do so causes duplicate IP addresses.
- After modifying the network, run the appropriate verification tools:
 - **For X8M and later, with RoCE Network Fabric:**

Run the `infinicheck` command to verify the network. You should supply a file that contains a comma-separated list of all the database server host names or RoCE Network Fabric IP addresses, and another file that lists all of the RoCE Network Fabric IP addresses for the storage servers. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts -c cells
```

```
INFINICHECK  
[Network Connectivity, Configuration and  
Performance]
```

```
#### FABRIC TYPE TESTS ####
```

```
System type identified: RoCE
```

```
Verifying User Equivalence of user=root from all DBs to all  
CELLs.
```

```
#### RoCE CONFIGURATION TESTS ####  
Checking for presence of RoCE devices on all DBs and  
CELLs
```

```
[SUCCESS].... RoCE devices on all DBs and CELLs look good
```

```
Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS].... RoCE Policy Routing settings look good

Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS].... RoCE DSCP ToS settings look good

Checking for RoCE PFC settings and DSCP mapping on all DBs
and CELLS
[SUCCESS].... RoCE PFC and DSCP settings look good

Checking for RoCE interface MTU settings. Expected value :
2300
[SUCCESS].... RoCE interface MTU settings look good

Verifying switch advertised DSCP on all DBs and CELLS ports
( ~ 2 min )
[SUCCESS].... Advertised DSCP settings from RoCE switch looks good

##### CONNECTIVITY TESTS #####
[COMPUTE NODES -> STORAGE CELLS]
(60 seconds approx.)
(Will walk through QoS values: 0-6)
[SUCCESS].....Results OK

[SUCCESS]..... All can talk to all storage cells

[COMPUTE NODES -> COMPUTE NODES]
(60 seconds approx.)
(Will walk through QoS values: 0-6)
[SUCCESS].....Results OK

[SUCCESS]..... All hosts can talk to all other nodes

Verifying Subnet Masks on all nodes
[SUCCESS] ..... Subnet Masks is same across the network
```

If user equivalence for password-less SSH is not configured, then you must first run `infinicheck` with the `-s` option. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts -c cells -s
```

- For X8 and earlier, with InfiniBand Network Fabric:

Run the `verify-topology` (or InfiniBand commands like `showtopology` and `ibdiagnet`) and `infinicheck` commands to verify the network is working properly. For example:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify-topology -t fattree
# ./infinicheck -g hosts -c cells
```

- When Exadata Database Machines run in separate clusters, do not modify the `cellip.ora` files. The `cellip.ora` file on a database server should only include the IP addresses for the storage servers used with that database server.
- Cells with disk types different from what is already installed can be added, but the disk types cannot be mixed in the same Oracle Automatic Storage Management (Oracle ASM) disk group. For example, if the existing disk groups all use high performance disks, and cells with high capacity disks are being added, then it is necessary to create new disk groups for the high capacity disks.

When adding the same type of disk, ensure that the grid disk sizes are exactly the same even if the new disks are larger than the existing ones. For example, if the existing disks are 3 TB, and the additional disks are 4 TB, then it is necessary to create grid disks that match the size on the 3 TB disks. A new disk group can be created using the extra 1 TB of disk space.

- In order to access Exadata Storage Servers in one Exadata Database Machine by another Exadata Database Machine when they are not running as a single cluster, Exadata Storage Servers must have unique Oracle ASM disk group and failure group names on each Exadata Database Machine.

For example, for two Exadata Database Machines cabled together but run as separate clusters, the following names should be unique:

- Cell name
 - Cell disk name
 - Grid disk name
 - Oracle ASM failure group name
- All equipment receives a Customer Support Identifier (CSI). Any new equipment for the Exadata Database Machine has a new CSI. Contact Oracle Support Services to reconcile the new CSI with the existing Exadata Database Machine CSI. Have the original instance numbers or serial numbers available, as well as the new numbers when contacting Oracle Support Services.

For X8M and later, with RoCE Network Fabric:

You can use the RDMA Network Fabric for limited external connectivity. The external connectivity ports in the RoCE Network Fabric switches can connect to Oracle ZFS Storage Appliance or Oracle Zero Data Loss Recovery Appliance to provide a backup solution.

On each leaf switch, use the available ports as described in [Exadata X8M Backup with ZFS Storage Appliance](#). For high availability connections, ensure that you spread the physical network connections from the storage appliance across both leaf switches.

The validated RoCE cable lengths are:

- Up to 5 meters for passive copper QSFP28 cables or QSFP28 to SFP28 breakout cables

- Up to 100 meters for fiber optic QSFP28 cables

For X8 and earlier, with InfiniBand Network Fabric:

The RDMA Network Fabric can be used for external connectivity. The external connectivity ports in the Sun Datacenter InfiniBand Switch 36 switches can connect to media servers for tape backup, data loading, and client and application access. Use the available ports on the leaf switches for external connectivity. There are 12 ports per rack. The available ports are 5B, 6A, 6B, 7A, 7B, and 12A in each leaf switch. For high availability connections, connect one port to one leaf switch and the other port to the second leaf switch. The validated InfiniBand cable lengths are:

- Up to 5 meters for passive copper 4X QDR QSFP cables
- Up to 100 meters for fiber optic 4X QDR QSFP cables

Related Topics

- Elastic Configurations

1.2 Reviewing the Safety Precautions

Before upgrading Oracle Exadata Database Machines, read *Important Safety Information for Sun Hardware Systems* included with the rack.

 **Note:**

Contact a service representative or Oracle Advanced Customer Support to confirm that Oracle has qualified your equipment for installation and use in Oracle Exadata Database Machine. Oracle is not liable for any issues when you install or use non-qualified equipment.

 **See Also:**

- *Oracle Exadata Database Machine Installation and Configuraton Guide* for safety guidelines
- *Oracle Engineered System Safety and Compliance Guide, Compliance Model No.: ESY27* for safety notices

1.3 Reviewing the Cable Precautions

Review the following RDMA Network Fabric cable precautions before working with the cables:

- Fiber optic cables with laser transceivers must be of type Class 1.
- Do not allow any copper core cable to bend to a radius tighter than 127 mm (5 inches). Tight bends can damage the cables internally.
- Do not allow any optical cable to bend to a radius tighter than 85 mm (3.4 inches). Tight bends can damage the cables internally.

- Do not use zip ties to bundle or support cables. The sharp edges of the ties can damage the cables internally. Use hook-and-loop straps.
- Do not allow any cable to experience extreme tension. Do not pull or drag the cables, which can damage them internally.
- Unroll a cable for its length.
- Do not twist cables more than one revolution for their entire length. Twisting a cable might damage it internally.
- Do not route cables where they can be stepped on, or experience rolling loads. A crushing effect can damage the cable internally.

1.4 Estimating Cable Path Lengths

Cable paths should be as short as possible. When the length of a cable path has been calculated, select the shortest cable to satisfy the length requirement. When specifying a cable, consider the following:

- Bends in the cable path increase the required length of the cable. Rarely does a cable travel in a straight line from connector to connector. Bends in the cable path are necessary, and each bend increases the total length.
- Bundling increases the required length of the cables. Bundling causes one or more cables to follow a common path. However, the bend radius is different in different parts of the bundle. If the bundle is large and unorganized, and there are many bends, one cable might experience only the inner radius of bends, while another cable might experience the outer radius of bends. In this situation, the differences of the required lengths of the cables is quite substantial.
- When calculating the cable path length and route is under the floor, take into consideration the height of the raised floor.

1.5 Bundling Cables

When bundling RDMA Network Fabric cables in groups, use hook-and-loop straps to keep cables organized. If possible, use color-coordinated straps to help identify cables and their routing. Splitter and C-4 copper conductor cables are fairly thick and heavy for their length. Consider the retention strength of the hook-and-loop straps when supporting cables. Bundle as few cables as reasonably possible. If the cables break free of their straps and fall free, the cables might break internally when they strike the floor or from sudden changes in tension.

Bundle the cables using many hook-and-loop straps. Oracle recommends that no more than eight cables be bundled together.

Place the hook-and-loop straps as close together as reasonably possible, for example, one strap every foot (0.3 m). If a cable breaks free from a strap, then the cable cannot fall far before it is retained by another strap.

- [Floor and Underfloor Delivery of Cables](#)

1.5.1 Floor and Underfloor Delivery of Cables

The RDMA Network Fabric switches accept cables from floor or underfloor delivery. Floor and underfloor delivery limits the tension in the cable to the weight of the cable for the rack height of the switch.



Note:

Overhead cabling details are not included in this guide. For details on overhead cabling, contact a certified service engineer.

1.6 Reviewing the Cable Management Arm Guidelines

Review the following cable management arm (CMA) guidelines before routing the cables.

- Remove all required cables from the packaging, and allow cables to acclimate or reach operating temperature, if possible. The acclimation period is usually 24 hours. This improves the ability to manipulate the cables.
- Label both ends of each cable using a label stock that meets the ANSI/TIA/EIA 606-A standard, if possible.
- Begin the installation procedure in ascending order.
- Only slide out one server at a time. Sliding out more than one server can cause cables to drop cause problems when sliding the servers back.
- Separate the installation by dressing cables with the least stringent bend radius requirements first. The following bend radius requirements are based on EIA/TIA 568-x standards, and may vary from the manufacturer's requirements:
 - CAT5e UTP: 4 x diameter of the cable or 1 inch; 25.4 mm minimum bend radius
 - AC power cables: 4 x diameter of the cable or 1 inch; 25.4 mm minimum bend radius
 - **For X8M and later, with RoCE Network Fabric:**
 - * 30 AWG: Single cable diameter of 4.5 +/- 0.2 mm and lengths from 1 to 3 meters; 21 mm single bend minimum bend radius or 45 mm repeated bends.
 - * 26 AWG: Single cable diameter of 5.8 +0.3 mm/-1.0 mm and lengths from 2.5 to 5 meters; 29 mm single bend minimum bend radius or 58 mm repeated bends.
 - **For X8 and earlier, with InfiniBand Network Fabric:**
 - * TwinAx: 5 x diameter of the cable or 1.175 inch; 33 mm minimum bend radius.
 - * Quad Small Form-factor Pluggable (QSFP) cable: 6 x diameter of the cable or 2 inch; 55 mm minimum bend radius.
 - * Fiber core cable: 10 x diameter of the cable or 1.22 inch; 31.75 mm minimum bend radius for a 0.125 cable.
- Install the cables with the best longevity rate first.

1.7 Obtaining Current Configuration Information

The current configuration information is used to plan patching requirements, configure new IP addresses, and so on. The following information should be collected as described before extending the rack:

- The Oracle EXAchk report for the current rack.
- Image history information using the following command:

```
dcli -g ~/all_group -l root "imagehistory" > imagehistory.txt
```

- Current IP addresses defined for all Exadata Storage Servers and database servers using the following command:

```
dcli -g ~/all_group -l root "ifconfig" > ifconfig_all.txt
```

- Information about the configuration of the cells, cell disks, flash logs, and IORM plans using the following commands:

```
dcli -g ~/cell_group -l root "cellcli -e list cell detail" >  
cell_detail.txt
```

```
dcli -g ~/cell_group -l root "cellcli -e list physicaldisk detail"  
> \  
physicaldisk_detail.txt
```

```
dcli -g ~/cell_group -l root "cellcli -e list griddisk  
attributes \  
name,offset,size,status,asmmodestatus,asmdeactivationoutcome" >  
griddisk.txt
```

```
dcli -g ~/cell_group -l root "cellcli -e list flashcache detail"  
> \  
fc_detail.txt
```

```
dcli -g ~/cell_group -l root "cellcli -e list flashlog detail" >  
fl_detail.txt
```

```
dcli -g ~/cell_group -l root "cellcli -e list iormplan detail"  
> \  
iorm_detail.txt
```

- HugePages memory configuration on the database servers using the following command:

```
dcli -g ~/dbs_group -l root "cat /proc/meminfo | grep 'HugePages'"  
> \  
hugepages.txt
```

- **For X8M and later, with RoCE Network Fabric:**

You should have a list of the switch names from the initial configuration with OEDA. Put the names in a file, for example, `roceswitches.txt`, with each switch name on a separate line.

- **For X8 and earlier, with InfiniBand Network Fabric:**

Use the following command:

```
ibswitches > ibswitches.txt
```

Use the `nm2version` on each Sun Datacenter InfiniBand Switch 36 switch to get its firmware version.

- The following network files from the first database server in the rack:

- `/etc/resolv.conf`
- `/etc/ntp.conf`
- `/etc/network`
- `/etc/sysconfig/network-scripts/ifcfg-*`

- Any users, user identifiers, groups and group identifiers created for cluster-managed services that need to be created on the new servers, such as Oracle GoldenGate.

- `/etc/passwd`
- `/etc/group`

- Output of current cluster status using the following command:

```
crsctl stat res -t > crs_stat.txt
```

- Patch information from the Grid Infrastructure and Oracle homes using the following commands. The commands must be run as Grid Infrastructure home owner, and the Oracle home owner.

```
/u01/app/oracle/product/11.2.0/dbhome_1/OPatch/opatch lsinventory -oh \
GRID_HOME -detail -all_nodes > opatch_grid.txt
```

```
/u01/app/oracle/product/11.2.0/dbhome_1/OPatch/opatch lsinventory -oh \
ORACLE_HOME -detail -all_nodes >> opatch_oracle.txt
```

In the preceding commands, `GRID_HOME` is the path for the Grid Infrastructure home directory, and `ORACLE_HOME` is the path for the Oracle home directory.

Related Topics

- [Oracle Autonomous Health Framework User's Guide](#)
- [Oracle Exadata Database Machine exachk or HealthCheck \(Doc ID 1070954.1\)](#)

1.8 Preparing the Network Configuration

When adding additional servers to your rack, you will need IP address and the current network configuration settings.

When adding additional servers or rack to an existing rack, the IP addresses for the new servers are obtained using Oracle Exadata Deployment Assistant (OEDA). If adding additional servers to an existing rack, then the application should only include the new servers. If adding an additional rack, then the new rack should use its own OEDA. The exact Oracle Automatic Storage Management (Oracle ASM) disk group configuration currently in use may not be reflected by the application. This is not an issue, as the grid disks and disk groups are configured manually. All other items, such as the Oracle home location and owner, should be defined exactly as the existing configuration.

When adding Oracle Exadata X4-2 Database Server or later or Oracle Exadata Storage Server X4-2L or later, the bonding configuration must match the existing servers in the rack. The OEDA InfiniBand configuration page has an option to select the type of bonding. Select the option for active-active bonding, or deselect the option for active-passive bonding.

If you want to configure network isolation for groups of virtual database servers and storage servers so that network traffic of one Oracle RAC cluster is not accessible to another Oracle RAC cluster, then you can use OEDA to configuration InfiniBand Partitioning.

The configuration file generated by the application is used by OEDA. After using OEDA, use the `checkip.sh` and `dbm.dat` files to verify the network configuration. The only errors that should occur are from the `ping` command to the SCAN addresses, Cisco switch, and Sun Datacenter InfiniBand Switch 36 switches.

Related Topics

- [Implementing InfiniBand Partitioning across OVM RAC Clusters on Oracle Exadata](#)
- [Cluster Networks - OEDA Online Help page](#)
- [Configuring New Servers Manually](#)
When adding servers to an Exadata Database Machine, you can configure the servers manually instead of using OEDA.

1.9 Moving Audit and Diagnostic Files

The files in the `$GRID_HOME/rdbms/audit` directory and the `$GRID_HOME/log/diagnostics` directory should be moved or deleted before extending a cluster. Oracle recommends moving or deleting the files a day or two before the planned extension because it may take time.

1.10 Reviewing Release and Patch Levels

When adding new servers to a rack, you must match the installed operating system version and software releases.

The new rack or servers most-likely have a later release or patch level than the current rack. In some cases, you may want to update the current rack release to the later release. In other cases, you may want to stay at your current release, and choose to re-image the new rack to match the current rack. Whatever you choose to do, ensure that the existing and new servers and switches are at the same patch level.

**Tip:**

Check My Oracle Support note 888828.1 for latest information on minimum releases.

Older servers in a rack may need to be patched to a later release to meet the minimum required software release. In addition, older database servers might use Oracle Linux release 5.3. Those servers need to be updated to a newer Oracle Linux release.

Additional patching considerations include the Oracle Grid Infrastructure and Oracle Database home releases and updates. If new patches will be applied, then Oracle recommends changing the existing servers so that the new servers will inherit the releases as part of the extension procedure. This way, the number of servers being patched is lower. Any patching of the existing servers should be performed in advance so they are at the desired level when the extension work is scheduled, thereby reducing the total amount of work required during the extension.

Related Topics

- [Exadata Database Machine and Exadata Storage Server Supported Versions \(My Oracle Support Doc ID 888828.1\)](#)
- [Updating key software components on database hosts to match those on the cells \(My Oracle Support Doc ID 1284070.1\)](#)

1.11 Performing Preliminary Checks

Perform a visual check of Oracle Exadata Database Machine physical systems before extending the hardware.

1. Check the rack for damage.
2. Check the rack for loose or missing screws.
3. Check Oracle Exadata Database Machine for the ordered configuration.
4. Check that all cable connections are secure and well seated.
5. Check power cables.
6. Ensure the correct connectors have been supplied for the data center facility power source.
7. Check network data cables.
8. Check the site location tile arrangement for cable access and airflow.
9. Check the data center airflow into the front of Oracle Exadata Database Machine.

1.12 Preparing to Add Servers

Perform the following tasks before adding the servers:

1. Unpack the Oracle Exadata Database Machine expansion kit.
2. Unpack all Oracle Exadata Database Machine server components from the packing cartons. The following items should be packaged with the servers:

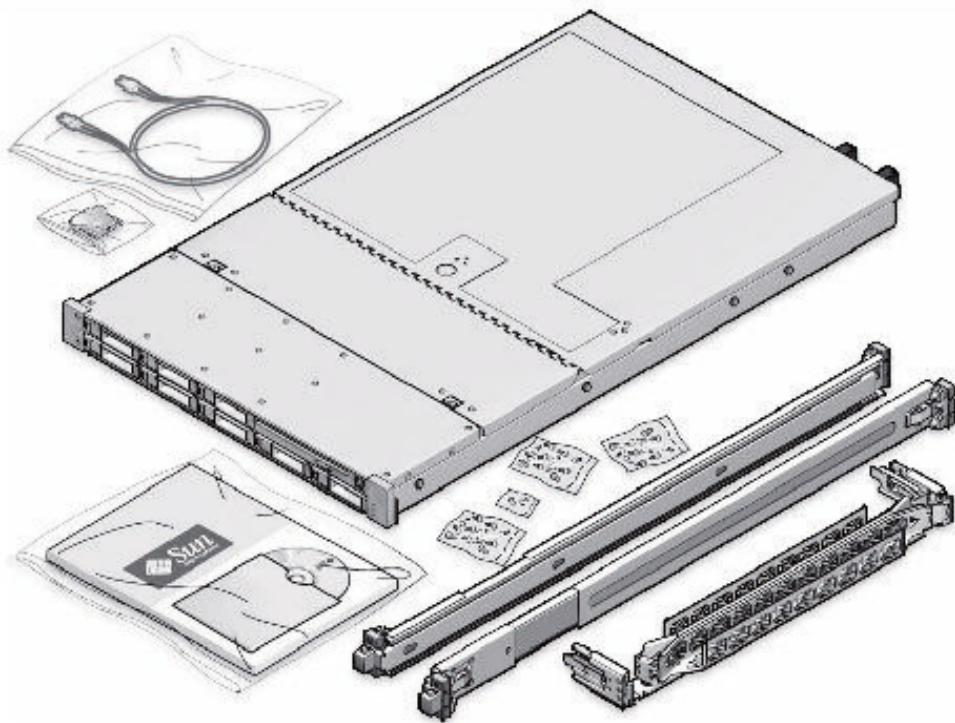
- Oracle Database servers or Exadata Storage Server
- Power cord, packaged with country kit
- Cable management arm with installation instructions
- Rackmount kit containing rack rails and installation instructions
- (Optional) Sun server documentation and media kit

 **Note:**

If you are extending Oracle Exadata Database Machine X4-2, Oracle Exadata Database Machine X3-8 Full Rack, or Oracle Exadata Database Machine X2-2 (with X4170 and X4275 servers) half rack, then order the expansion kit that includes a Sun Datacenter InfiniBand Switch 36 switch.

Figure 1-1 shows the components in the server expansion kit.

Figure 1-1 Server Components for Upgrade



3. Lay out the cables for the servers.
4. Unroll the cables and stretch them to remove the bends.
5. Apply the cable labels. Oracle recommends labeling all cables before installation.
6. Install the servers.
7. Cable the servers.

 **See Also:**

- *Oracle Exadata Database Machine Maintenance Guide* for information about cable labels
- "[Adding New Servers](#)" for information about installing the servers
- "[Cabling Database Servers](#)" and "[Cabling Exadata Storage Servers](#)" for information about cabling the servers

2

Extending the Hardware

You can extend Oracle Exadata Database Machine by adding database servers and storage servers within a rack. You can also cable together multiple racks.

All new equipment receives a Customer Support Identifier (CSI). Any new equipment for your Oracle Exadata Rack has a new CSI. Contact Oracle Support Services to reconcile the new CSI with the existing Oracle Exadata Rack CSI. Have the original instance numbers or serial numbers available, as well as the new numbers when contacting Oracle Support Services.

- [Extending an Eighth Rack](#)
- [Extending Elastic Configurations](#)
Exadata Database Machine is available in Elastic Configurations that consist of a number of database and storage servers up to the capacity of the rack, as defined within Oracle Exadata Configuration Assistant (OECA).
- [Extending a Rack by Adding Another Rack](#)
You can extend your Oracle Exadata Rack by adding another rack and configuring the racks together.

2.1 Extending an Eighth Rack

The process of extending an Oracle Exadata Database Machine Eighth Rack system is principally the same as extending any other Exadata rack. At a high level, the process involves the following stages:

1. Verify the preexisting Eighth Rack system configuration.
2. Perform any required hardware installation.

During this stage, you may need to shut down servers to install some hardware components. Where required, servers would be shut down in a rolling manner to preserve overall system availability.

3. Configure the additional hardware resources.

During this stage:

- New database server resources are added to a database cluster.
- New storage server resources are used to create additional grid disks, which can be used to extend existing Oracle ASM disk groups.

The following sections contain additional information and procedures specific to Oracle Exadata Database Machine Eighth Rack:

- [Eighth Rack Expansion Options](#)
This section outlines the available expansion options for different Oracle Exadata Database Machine Eighth Rack models.
- [Eighth Rack Expansion Procedures](#)
This section contains specific procedures that may be required for expanding some Oracle Exadata Database Machine Eighth Rack models.

2.1.1 Eighth Rack Expansion Options

This section outlines the available expansion options for different Oracle Exadata Database Machine Eighth Rack models.

- [X9M-2 Eighth Rack Upgrade Options](#)
- [X7-2, X8-2, and X8M-2 Eighth Rack Upgrade Options](#)
- [X6M-2 Eighth Rack Upgrade Options](#)
- [Eighth Rack X4-2 and X5-2 Expansion Options](#)

2.1.1.1 X9M-2 Eighth Rack Upgrade Options

Upgrade of Oracle Exadata Database Machine X9M-2 Eighth Rack systems requires hardware modification. Compared with non-Eighth Rack systems, Eighth Rack database servers have one CPU, and less memory, and Eighth Rack storage servers have fewer CPU cores, less memory, less disk storage, and less flash storage.

You can upgrade Oracle Exadata Database Machine X9M-2 Eighth Rack systems as follows:

- Add a CPU upgrade kit to the database servers, which adds 1 CPU to each database server.
As part of adding the CPU upgrade kit, you must also:
 1. Physically move the network interface card from PCIe slot 2 to PCIe slot 1.
 2. Run `/opt/oracle.cellos/ipconf.pl` to reconfigure the OS network configuration files.
- Add a memory upgrade kit to the database servers, which adds 4 x 32 GB DIMMs to each database server.

On database servers with 384 GB of memory, you must install the memory upgrade kit if you want to install the CPU upgrade kit.

- Add more database servers.

You can expand processing capacity by adding standard (fully-configured) database servers. Though not mandatory, upgrading the original Eighth Rack database servers before adding more database servers is recommended to maximize consistency across the rack.

- Add more storage servers.

There is no upgrade kit to expand Eighth Rack High Capacity storage servers. However, you can expand storage capacity by adding more Eighth Rack High Capacity storage servers.

You can also expand storage capacity by adding standard High Capacity (HC), Extreme Flash (EF), or Extended (XT) storage servers.

2.1.1.2 X7-2, X8-2, and X8M-2 Eighth Rack Upgrade Options

Upgrade of Oracle Exadata Database Machine X7-2, X8-2, or X8M-2 Eighth Rack systems requires hardware modification. Eighth Rack database servers have one of the CPUs removed, and all of the memory for CPU1 was moved to CPU0. Storage servers have half the cores enabled, and half the disks and flash cards were removed.

On Oracle Exadata Database Machine X7-2, X8-2, or X8M-2 Eighth Rack systems with Extreme Flash storage servers, you can add CPUs and flash cards to extend the system to a Quarter Rack.

For Oracle Exadata Database Machine X7-2, X8-2, or X8M-2 Eighth Rack systems with High Capacity storage servers, you can add the CPU and memory to the database servers and additional Eighth Rack High Capacity storage servers to expand the system.

Specifically:

- To extend an Eighth Rack X7-2, X8-2, or X8M-2 database server:
 1. Install CPU1
 2. Move half of CPU0's memory to CPU1
 3. Move the 10/25GbE PCI card to PCIe slot 1
- To extend an Eighth Rack X7-2, X8-2, or X8M-2 Extreme Flash storage server, install four F640/F640v2 flash cards in PCIe slots 2,3,8, and 9.

2.1.1.3 X6M-2 Eighth Rack Upgrade Options

On Oracle Exadata Database Machine X6-2 Eighth Rack systems, you can add High Capacity (HC) storage by adding more Eighth Rack High Capacity storage servers. There is no upgrade kit to expand individual Eighth Rack High Capacity storage servers.

Additionally, software-only upgrade kits are available to enable the disabled hardware resources for X6-2 Eighth Rack database servers and X6-2 Eighth Rack Extreme Flash (EF) storage servers.

2.1.1.4 Eighth Rack X4-2 and X5-2 Expansion Options

Oracle Exadata Database Machine Eighth Rack X4-2 and X5-2 contains fully configured servers, with approximately half of the available system resources (CPU, RAM, and storage) disabled. Consequently, Eighth Rack to Quarter Rack expansion of X4-2 or X5-2 systems is done using software only to enable the disabled hardware resources. Expansion beyond a Quarter Rack requires additional expansion kits, which contain additional hardware.

2.1.2 Eighth Rack Expansion Procedures

This section contains specific procedures that may be required for expanding some Oracle Exadata Database Machine Eighth Rack models.

Some of the following procedures only apply in specific circumstances. So, if present, check the applicability note at the beginning of the procedure. Otherwise, the expansion of an Eighth Rack through the addition of database or storage servers uses the same general procedures as the elastic expansion of a non-Eighth Rack. See [Extending Elastic Configurations](#) and [Configuring the New Hardware](#).

Also, in the following procedures:

- The disk group names and sizes are examples. The values should be changed in the commands to match the actual system.
- User equivalence (password-less SSH) is assumed to exist between the `root` user on the first database server and all other database servers, and to the `celladmin` user on all storage cells.

- The text files `cell_group` and `db_group` should be created to contain lists of cell host names and database server host names, respectively.

Eighth Rack Expansion Procedures:

- [Reviewing and Validating Current Configuration of Eighth Rack Oracle Exadata Database Machine](#)
The following procedure describes how to review and validate the current configuration.
- [Activating Database Server Cores in Oracle Exadata Database Machine Eighth Rack](#)
The following procedure describes how to activate database server cores.
- [Activating Storage Server Cores and Disks in Oracle Exadata Database Machine Eighth Rack](#)
The following procedure describes how to activate the storage server cores and disks.
- [Creating Additional Grid Disks in Oracle Exadata Database Machine Eighth Rack](#)
Additional grid disk creation must follow a specific order to ensure the proper offset.
- [Adding Grid Disks to Oracle ASM Disk Groups in Oracle Exadata Database Machine Eighth Rack](#)
The following procedure describes how to add the new grid disks to Oracle ASM disk groups.
- [Validating Expansion of Oracle Exadata Database Machine](#)
After expansion, validate the new configuration.

2.1.2.1 Reviewing and Validating Current Configuration of Eighth Rack Oracle Exadata Database Machine

The following procedure describes how to review and validate the current configuration.

1. Log in as the `root` user on the first database server.
2. Review the current configuration of the storage servers using the following command. The expected output is `TRUE`.

```
# dcli -g cell_group -l celladmin 'cellcli -e LIST CELL attributes eighthrack'
```

3. Review the current CPU core count on the database servers using the following command:

```
# dcli -g db_group -l root 'dbmcli -e list dbserver attributes coreCount'
```

The following is an example of the expected output from an Oracle Exadata Database Machine X9M-2 Eighth Rack with all CPU cores enabled:

```
dm01db01: 32  
dm01db02: 32
```

Contact Oracle Support Services if the number of active database server CPU cores differs from the expected value.

Related Topics

- Oracle Exadata Database Server Hardware Components

2.1.2.2 Activating Database Server Cores in Oracle Exadata Database Machine Eighth Rack

The following procedure describes how to activate database server cores.

Note:

This procedure applies to:

- Original database server CPU cores that are disabled
- Additional CPU cores that are part of an approved CPU upgrade kit

The following is not required where database server hardware resources are expanded by adding more database servers.

1. Log in as the `root` user on the first database server.
2. Activate the database server cores using the following `dcli` utility command on the database server group:

```
# dcli -g db_group -l root 'dbmcli -e \
ALTER DBSERVER pendingCoreCount = number_of_cores'
```

In the preceding command, *number_of_cores* is the total number of cores to activate. The value includes the existing core count and the additional cores to be activated. The following command shows how to activate all the cores in Oracle Exadata Database Machine X5-2 Eighth Rack:

```
# dcli -g db_group -l root 'dbmcli -e ALTER DBSERVER pendingCoreCount =
36'
```

For a description of the supported core counts for each server model, see [Restrictions for Capacity-On-Demand on Oracle Exadata Database Machine](#)

3. Restart each database server.

 **Note:**

If this procedure is done in a rolling fashion with the Oracle Database and Oracle Grid Infrastructure active, then ensure the following before restarting the database server:

- All Oracle ASM grid disks are online.
- There are no active Oracle ASM rebalance operations. You can query the `V$ASM_OPERATION` view for the status of the rebalance operation.
- Shut down Oracle Database and Oracle Grid Infrastructure in a controlled manner, failing over services as needed.

4. Verify the following items on the database server after the restart completes and before proceeding to the next server:
 - The Oracle Database and Oracle Grid Infrastructure services are active.
See *Using SRVCTL to Verify That Instances are Running in Oracle Real Application Clusters Administration and Deployment Guide* and the `crsctl status resource -w "TARGET = ONLINE" -t` command.
 - The number of active cores is correct. Use the `dbmcli -e list dbserver attributes coreCount` command to verify the number of cores.

 **See Also:**

- Changing a Disk to Offline or Online
- Stopping One or More Instances and Oracle RAC Databases Using SRVCTL in *Oracle Real Application Clusters Administration and Deployment Guide*
- `crsctl stop cluster`
- Increasing the Number of Active Cores on Database Servers

2.1.2.3 Activating Storage Server Cores and Disks in Oracle Exadata Database Machine Eighth Rack

The following procedure describes how to activate the storage server cores and disks.

 **Note:**

This procedure applies only to the original storage servers in the following Oracle Exadata Database Machine Eighth Rack models: X4-2, X5-2, and X6-2 with Extreme Flash (EF) storage servers.

This procedure does not apply where storage server hardware resources are expanded by adding more storage servers.

1. Log in as the `root` user on the first database server.
2. Activate the cores on the storage server group using the following command. The command uses the `dcli` utility, and runs the command as the `celladmin` user.

```
# dcli -g cell_group -l celladmin cellcli -e "alter cell eighthRack=false"
```

3. Create the cell disks using the following command:

```
# dcli -g cell_group -l celladmin cellcli -e "create celldisk all"
```

4. Recreate the flash log using the following commands:

```
# dcli -g cell_group -l celladmin cellcli -e "drop flashlog all force"  
# dcli -g cell_group -l celladmin cellcli -e "create flashlog all"
```

5. Expand the flash cache using the following command:

```
# dcli -g cell_group -l celladmin cellcli -e "alter flashcache all"
```

2.1.2.4 Creating Additional Grid Disks in Oracle Exadata Database Machine Eighth Rack

Additional grid disk creation must follow a specific order to ensure the proper offset.

 **Note:**

This procedure applies only to the original storage servers in the following Oracle Exadata Database Machine Eighth Rack models: X4-2, X5-2, and X6-2 with Extreme Flash (EF) storage servers.

This procedure does not apply where storage server hardware resources are expanded by adding more storage servers.

The order of grid disk creation must follow the same sequence that was used during the initial grid disk creation process. For a standard deployment using Oracle Exadata Deployment Assistant (OEDA), the order is DATA, RECO, and DBFS_DG (if present). Create all DATA grid disks first, followed by the RECO grid disks, and then the DBFS_DG grid disks (if present).

The following procedure describes how to create the grid disks:

 **Note:**

The commands shown in this procedure use the standard deployment grid disk prefix names of DATA, RECO, and DBFS_DG. The sizes being checked are on cell disk 02. Cell disk 02 is used because the disk layout for cell disks 00 and 01 are different from the other cell disks in the server.

1. Check the size of the grid disks using the following commands. Each cell should return the same size for the grid disks starting with the same grid disk prefix.

```
# dcli -g cell_group -l celladmin cellcli -e \
"list griddisk attributes name, size where name like
\'DATA.*_02_.*\'"
```

```
# dcli -g cell_group -l celladmin cellcli -e \
"list griddisk attributes name, size where name like
\'RECO.*_02_.*\'"
```

```
# dcli -g cell_group -l celladmin cellcli -e \
"list griddisk attributes name, size where name like
\'DBFS_DG.*_02_.*\'"
```

The sizes shown are used during grid disk creation.

2. Create the grid disks for the disk groups using the sizes shown in step 1. The following table shows the commands to create the grid disks based on rack type and disk group.

Table 2-1 Commands to Create Disk Groups When Extending Oracle Exadata Database Machine Eighth Rack

Rack	Commands
Extreme Flash Oracle Exadata Database Machine	<pre> dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_FD_04_\'hostname -s\' celldisk=FD_04_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_FD_05_\'hostname -s\' celldisk=FD_05_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_FD_06_\'hostname -s\' celldisk=FD_06_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_FD_07_\'hostname -s\' celldisk=FD_07_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_FD_04_\'hostname -s\' celldisk=FD_04_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_FD_05_\'hostname -s\' celldisk=FD_05_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_FD_06_\'hostname -s\' celldisk=FD_06_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_FD_07_\'hostname -s\' celldisk=FD_07_\'hostname - </pre>

Table 2-1 (Cont.) Commands to Create Disk Groups When Extending Oracle Exadata Database Machine Eighth Rack

Rack	Commands
	<pre> s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_FD_04_\'hostname -s\' celldisk=FD_04_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_FD_05_\'hostname -s\' celldisk=FD_05_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_FD_06_\'hostname -s\' celldisk=FD_06_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_FD_07_\'hostname -s\' celldisk=FD_07_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" </pre>

Table 2-1 (Cont.) Commands to Create Disk Groups When Extending Oracle Exadata Database Machine Eighth Rack

Rack	Commands
High Capacity Oracle Exadata Database Machine	<pre> dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_CD_06_\'hostname -s\' celldisk=CD_06_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_CD_07_\'hostname -s\' celldisk=CD_07_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_CD_08_\'hostname -s\' celldisk=CD_08_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_CD_09_\'hostname -s\' celldisk=CD_09_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_CD_10_\'hostname -s\' celldisk=CD_10_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DATA_CD_11_\'hostname -s\' celldisk=CD_11_\'hostname - s\',size=datasize" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_CD_06_\'hostname -s\' celldisk=CD_06_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_CD_07_\'hostname -s\' celldisk=CD_07_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_CD_08_\'hostname -s\' celldisk=CD_08_\'hostname - s\',size=recoresize, \ </pre>

Table 2-1 (Cont.) Commands to Create Disk Groups When Extending Oracle Exadata Database Machine Eighth Rack

Rack	Commands
	<pre> cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_CD_09_\'hostname -s\' celldisk=CD_09_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_CD_10_\'hostname -s\' celldisk=CD_10_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ RECO_CD_11_\'hostname -s\' celldisk=CD_11_\'hostname - s\',size=recoresize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_CD_06_\'hostname -s\' celldisk=CD_06_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_CD_07_\'hostname -s\' celldisk=CD_07_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_CD_08_\'hostname -s\' celldisk=CD_08_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_CD_09_\'hostname -s\' celldisk=CD_09_\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ </pre>

Table 2-1 (Cont.) Commands to Create Disk Groups When Extending Oracle Exadata Database Machine Eighth Rack

Rack	Commands
	<pre>DBFS_DG_CD_10 \\'hostname -s\' celldisk=CD_10 \\'hostname - s\',size=dbfssize, \ cachingPolicy=none" dcli -g cell_group -l celladmin "cellcli -e create griddisk \ DBFS_DG_CD_11 \\'hostname -s\' celldisk=CD_11 \\'hostname - s\',size=dbfssize, \ cachingPolicy=none"</pre>

2.1.2.5 Adding Grid Disks to Oracle ASM Disk Groups in Oracle Exadata Database Machine Eighth Rack

The following procedure describes how to add the new grid disks to Oracle ASM disk groups.



Note:

This procedure applies only to the original storage servers in the following Oracle Exadata Database Machine Eighth Rack models: X4-2, X5-2, and X6-2 with Extreme Flash (EF) storage servers.

This procedure does not apply where storage server hardware resources are expanded by adding more storage servers.

The grid disks created in [Creating Additional Grid Disks in Oracle Exadata Database Machine Eighth Rack](#) must be added as Oracle ASM disks to their corresponding, existing Oracle ASM disk groups.

1. Validate the following:
 - No rebalance operation is currently running.
 - All Oracle ASM disks are active.
2. Log in to the first database server as the owner who runs the Oracle Grid Infrastructure software.
3. Set the environment to access the +ASM instance on the server.
4. Log in to the ASM instance as the `sysasm` user using the following command:

```
$ sqlplus / as sysasm
```

5. Validate the current settings, as follows:

```
SQL> set lines 100
SQL> column attribute format a20
SQL> column value format a20
SQL> column diskgroup format a20
SQL> SELECT att.name attribute, upper(att.value) value, dg.name
diskgroup
FROM V$ASM_ATTRIBUTE att, V$ASM_DISKGROUP DG
WHERE DG.group_number=att.group_number AND att.name LIKE
'%appliance.mode%'
ORDER BY att.group_number;
```

The output should be similar to the following:

ATTRIBUTE	VALUE	DISKGROUP
appliance.mode	TRUE	DATA1
appliance.mode	TRUE	DBFS_DG
appliance.mode	TRUE	RECO1

6. Disable the `appliance.mode` attribute for any disk group that shows `TRUE` using the following commands:

```
SQL> ALTER DISKGROUP data_diskgroup set attribute
'appliance.mode'='FALSE';
SQL> ALTER DISKGROUP reco_diskgroup set attribute
'appliance.mode'='FALSE';
SQL> ALTER DISKGROUP dbfs_dg_diskgroup set attribute
'appliance.mode'='FALSE';
```

In the preceding commands, `data_diskgroup`, `reco_diskgroup`, and `dbfs_dg_diskgroup` are the names of the DATA, RECO, and DBFS_DG disk groups, respectively.

7. Add the grid disks to the Oracle ASM disk groups. The following table shows the commands to create the grid disks based on rack type and disk group. Adding the new disks requires a rebalance of the system.

Table 2-2 Commands to Add Disk Groups When Extending Eighth Rack Oracle Exadata Database Machine

Rack	Commands
Extreme Flash Oracle Exadata Database Machine	<pre>SQL> ALTER DISKGROUP data_diskgroup ADD DISK 'o/*/ DATA_FD_0[4-7]*' \ REBALANCE POWER 32; SQL> ALTER DISKGROUP reco_diskgroup ADD DISK 'o/*/ RECO_FD_0[4-7]*' \ REBALANCE POWER 32; SQL> ALTER DISKGROUP dbfs_dg_diskgroup ADD DISK 'o/*/ DBFS_DG_FD_0[4-7]*'\ REBALANCE POWER 32;</pre>
High Capacity Oracle Exadata Database Machine	<pre>SQL> ALTER DISKGROUP data_diskgroup ADD DISK 'o/*/ DATA_CD_0[6-9]*',' \ o/*/DATA_CD_1[0-1]*' REBALANCE POWER 32; SQL> ALTER DISKGROUP reco_diskgroup ADD DISK 'o/*/ RECO_CD_0[6-9]*',' \ o/*/RECO_CD_1[0-1]*' REBALANCE POWER 32; SQL> ALTER DISKGROUP dbfs_dg_diskgroup ADD DISK ' \ o/*/DBFS_DG_CD_0[6-9]*',' o/*/DBFS_DG_CD_1[0-1]*' REBALANCE POWER 32;</pre>

The preceding commands return `Diskgroup altered`, if successful.

- (Optional) Monitor the current rebalance operation using the following command:

```
SQL> SELECT * FROM gv$asm_operation;
```

- Re-enable the `appliance.mode` attribute, if it was disabled in step 6 using the following commands:

```
SQL> ALTER DISKGROUP data_diskgroup set attribute 'appliance.mode'='TRUE';
SQL> ALTER DISKGROUP reco_diskgroup set attribute 'appliance.mode'='TRUE';
SQL> ALTER DISKGROUP dbfs_dg_diskgroup set attribute
'appliance.mode'='TRUE';
```

2.1.2.6 Validating Expansion of Oracle Exadata Database Machine

After expansion, validate the new configuration.

 **Note:**

This procedure applies only to the original storage servers in the following Oracle Exadata Database Machine Eighth Rack models: X4-2, X5-2, X6-2 with Extreme Flash (EF) storage servers.

This procedure does not apply where hardware resources are expanded by adding more servers.

1. Log in as the `root` user on the first database server.
2. Check the database server core count using the following command:

```
# dcli -g db_group -l root 'dbmcli -e list dbserver attributes
coreCount'
```

3. Review the database server configuration using the following command.

```
# dcli -g db_group -l root 'dbmcli -e list dbserver attributes
eighthtrack'
```

The output should show `FALSE`.

4. Review the storage server configuration using the following command.

```
# dcli -g cell_group -l celladmin 'cellcli -e list cell attributes
eighthtrack'
```

The output should show `FALSE`.

5. Review the appliance mode for each disk group using the following commands:

```
SQL> set lines 100
SQL> column attribute format a20
SQL> column value format a20
SQL> column diskgroup format a20
SQL> SELECT att.name attribute, upper(att.value) value, dg.name
diskgroup \
FROM V$ASM_ATTRIBUTE att, V$ASM_DISKGROUP
DG \
WHERE DG.group_number = att.group_number
AND \
att.name LIKE '%appliance.mode%' ORDER BY DG.group_number;
```

6. Validate the number of Oracle ASM disks using the following command:

```
SQL> SELECT
g.name,d.failgroup,d.mode_status,count(*) \
FROM v$asm_diskgroup g, v$asm_disk \
d \
WHERE
d.group_number=g.group_number
```

```
\
GROUP BY g.name,d.failgroup,d.mode_status;
```

NAME	FAILGROUP	MODE_ST	COUNT (*)
DATA1	EXA01CELADM01	ONLINE	12
DATA1	EXA01CELADM02	ONLINE	12
DATA1	EXA01CELADM03	ONLINE	12
RECOC1	EXA01CELADM01	ONLINE	12
RECOC1	EXA01CELADM02	ONLINE	12
RECOC1	EXA01CELADM03	ONLINE	12
RECOC2	EXA01CELADM01	ONLINE	12
RECOC2	EXA01CELADM02	ONLINE	12
RECOC2	EXA01CELADM03	ONLINE	12
DBFS_DG	EXA01CELADM01	ONLINE	10
DBFS_DG	EXA01CELADM02	ONLINE	10
DBFS_DG	EXA01CELADM03	ONLINE	10

Each High Capacity (HC) storage server (non-Eighth Rack) contains 12 disks.

2.2 Extending Elastic Configurations

Exadata Database Machine is available in Elastic Configurations that consist of a number of database and storage servers up to the capacity of the rack, as defined within Oracle Exadata Configuration Assistant (OECA).

Additional database and storage servers can be added if space is available; see OECA for details. The upgrade process includes adding new servers and cables



Note:

It is possible to extend the hardware while the machine is online, and with no downtime. However, extreme care should be taken. In addition, patch application to existing switches and servers should be done before extending the hardware.

- [Removing the Doors](#)
This procedure describes how to remove the doors on Exadata Database Machine.
- [Adding New RDMA Network Fabric Switches](#)
You can add individual new RDMA Network Fabric switches as needed to meet growing resource requirements.
- [Adding New Servers](#)
You can add new servers to an Oracle Exadata Rack that is not at full capacity.
- [Cabling Database Servers](#)
- [Cabling Storage Servers](#)
After the new Storage Servers are installed, you need to connect them to the existing equipment.
- [Closing the Rack](#)
After installing new equipment, you must replace the panels and close the rack.

2.2.1 Removing the Doors

This procedure describes how to remove the doors on Exadata Database Machine.

- If you are adding hardware to an Exadata Database Machine X7 rack, refer to "Remove the Doors" in *Oracle Rack Cabinet 1242 User's Guide* at https://docs.oracle.com/cd/E85660_01/html/E87280/gshfw.html#scrolltoc.
- If you are adding hardware to an earlier Exadata Database Machine rack, then refer to "Remove the Doors" in *Sun Rack II User's Guide* at <https://docs.oracle.com/cd/E19657-01/html/E29153/z40004911007624.html#scrolltoc>.

2.2.2 Adding New RDMA Network Fabric Switches

You can add individual new RDMA Network Fabric switches as needed to meet growing resource requirements.

The instructions are different for RoCE Network Fabric switches and InfiniBand Network Fabric switches.

- [Adding a RoCE Network Fabric Switch \(Cisco Nexus 9336C-FX2\)](#)
- [Adding an InfiniBand Network Fabric Switch \(Sun Datacenter InfiniBand Switch 36\)](#)

2.2.2.1 Adding a RoCE Network Fabric Switch (Cisco Nexus 9336C-FX2)

This procedure only applies to systems with Cisco Nexus 9336C-FX2 RoCE Network Fabric switches.



Note:

The steps in this procedure are specific to Exadata Database Machine. They are not the same as the steps in the Cisco Nexus manual.

1. Unpack the Cisco Nexus switch components from the packing cartons. The following items should be in the packing cartons:
 - Cisco Nexus 9336C-FX2 Switch
 - Cable bracket and rack-mount kit
 - Cable management bracket and cover
 - Two rack rail assemblies
 - Assortment of screws and captive nuts
 - Cisco Nexus 9336C-FX2 Switch documentationThe service label procedure on top of the switch includes descriptions of the preceding items.
2. Remove the trough from the rack in RU1. Put the cables aside while installing the RoCE Network Fabric switch. The trough can be discarded.
3. Install cage nuts in each rack rail in the appropriate holes.

4. Attach the brackets with cutouts to the power supply side of the switch.
5. Attach the C-brackets to the switch on the side of the ports.
6. Slide the switch halfway into the rack from the front. Keep the switch to the left side of the rack as far as possible while pulling the two power cords through the C-bracket on the right side.
7. Slide the server in rack location U2 out to the locked service position. This improves access to the rear of the switch during further assembly.
8. Install the slide rails from the rear of the rack into the C-brackets on the switch, pushing them up to the rack rail.
9. Attach an assembled cable arm bracket to the slide rail and using a No. 3 Phillips screwdriver, screw these together into the rack rail:
 - a. Install the lower screw loosely with the cable arm bracket rotated 90 degrees downward. This allows better finger access to the screw.
 - b. Rotate the cable arm bracket to the correct position.
 - c. Install the upper screw.
 - d. Tighten both screws.

If available, a screwdriver with a long-shaft (16-inch / 400mm) will allow easier installation such that the handle is outside the rack and beyond the cabling.
10. Push the switch completely into the rack from the front, routing the power cords through the cutout on the rail bracket.
11. Secure the switch to the front rack rail with M6 16mm screws. Tighten the screws using the No. 3 Phillips screwdriver.
12. Install the lower part of the cable management arm across the back of the switch.
13. Connect the cables to the appropriate ports.
14. Install the upper part of the cable management arm.
15. Slide the server in rack location U2 back into the rack.
16. Install power cords to the switch power supply slots on the front.
17. Loosen the front screws to install the vented filler panel brackets. Tighten the screws, and snap on the vented filler panel in front of the switch.

 **See Also:**

- *Oracle Exadata Database Machine System Overview* to view the rack layout
- *Oracle Exadata Database Machine System Overview* for information about networking cables

2.2.2.2 Adding an InfiniBand Network Fabric Switch (Sun Datacenter InfiniBand Switch 36)

This procedure only applies to systems with InfiniBand Network Fabric.

 **Note:**

The steps in this procedure are specific to Oracle Exadata Database Machine. They are not the same as the steps in the Sun Datacenter InfiniBand Switch 36 manual.

1. Unpack the Sun Datacenter InfiniBand Switch 36 switch components from the packing cartons. The following items should be in the packing cartons:
 - Sun Datacenter InfiniBand Switch 36 switch
 - Cable bracket and rackmount kit
 - Cable management bracket and cover
 - Two rack rail assemblies
 - Assortment of screws and captive nuts
 - Sun Datacenter InfiniBand Switch 36 documentation

The service label procedure on top of the switch includes descriptions of the preceding items.
2. X5 racks only: Remove the trough from the rack in RU1 and put the cables aside while installing the Sun Datacenter InfiniBand Switch 36 switch. The trough can be discarded.
3. Install cage nuts in each rack rail in the appropriate holes.
4. Attach the brackets with cutouts to the power supply side of the switch.
5. Attach the C-brackets to the switch on the side of the Sun Datacenter InfiniBand Switch 36 ports.
6. Slide the switch halfway into the rack from the front. You need to keep it to the left side of the rack as far as possible while pulling the two power cords through the C-bracket on the right side.
7. Slide the server in rack location U2 out to the locked service position. This improves access to the rear of the switch during further assembly.
8. Install the slide rails from the rear of the rack into the C-brackets on the switch, pushing them up to the rack rail.
9. Attach an assembled cable arm bracket to the slide rail and using a No. 3 Phillips screwdriver, screw these together into the rack rail:
 - a. Install the lower screw loosely with the cable arm bracket rotated 90 degrees downward. This allows better finger access to the screw.
 - b. Rotate the cable arm bracket to the correct position.
 - c. Install the upper screw.
 - d. Tighten both screws.

If available, a screwdriver with a long-shaft (16-inch / 400mm) will allow easier installation such that the handle is outside the rack and beyond the cabling.
10. Push the switch completely into the rack from the front, routing the power cords through the cutout on the rail bracket.

11. Secure the switch to the front rack rail with M6 16mm screws. Tighten the screws using the No. 3 Phillips screwdriver.
12. Install the lower part of the cable management arm across the back of the switch.
13. Connect the cables to the appropriate ports.
14. Install the upper part of the cable management arm.
15. Slide the server in rack location U2 back into the rack.
16. Install power cords to the Sun Datacenter InfiniBand Switch 36 switch power supply slots on the front.
17. Loosen the front screws to install the vented filler panel brackets. Tighten the screws, and snap on the vented filler panel in front of the switch.

 **See Also:**

- *Oracle Exadata Database Machine System Overview* to view the rack layout
- *Oracle Exadata Database Machine System Overview* for information about InfiniBand networking cables

2.2.3 Adding New Servers

You can add new servers to an Oracle Exadata Rack that is not at full capacity.

You can add individual database servers or storage servers to meet growing resource requirements using the Elastic Configuration method. See Oracle Exadata Configuration Assistant (OECA) for details. The upgrade process includes adding new servers and cables. Additional hardware may be required.

 **Note:**

- Always load equipment into the rack from the bottom up, so that the rack does not become top-heavy and tip over. Extend the rack anti-tip bar to prevent the rack from tipping during equipment installation.
- The new servers need to be configured manually.

- [Preparing to Install New Servers](#)
Before you install a new server, prepare the rack unit for the server installation.
- [Installing the Rack Assembly](#)
After preparing for installation, you next install the rack assembly to hold the new servers.
- [Installing the Server](#)
After preparing for the installation and installing the rack assembly, you then install the new server.

Related Topics

- [Setting Up New Servers](#)
New servers need to be configured when extending Exadata Database Machine Elastic Configurations.

2.2.3.1 Preparing to Install New Servers

Before you install a new server, prepare the rack unit for the server installation.

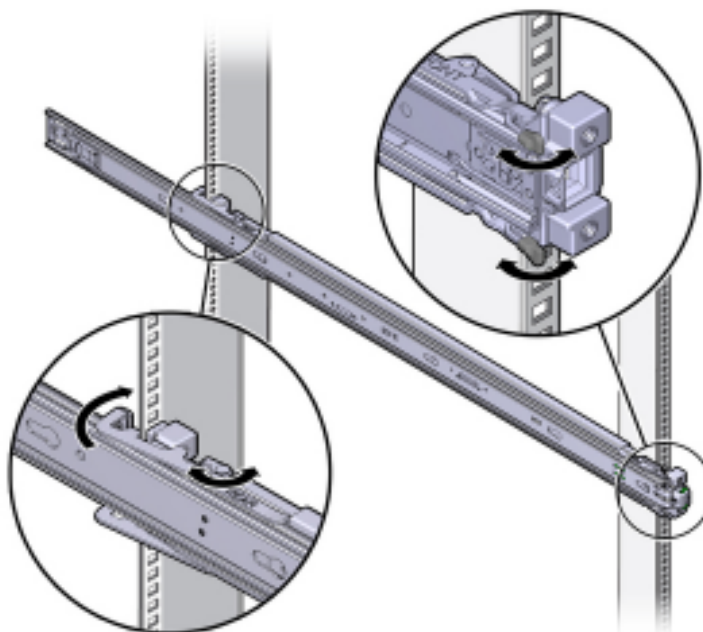
1. Identify the rack unit where the server will be installed. Fill the first available unit, starting from the bottom of the rack.
2. Remove and discard the trough, which attaches the cable harness when no server is installed in the unit.
3. Remove and discard the solid filler.

2.2.3.2 Installing the Rack Assembly

After preparing for installation, you next install the rack assembly to hold the new servers.

1. Position a mounting bracket against the chassis so that the slide-rail lock is at the server front, and the five keyhole openings on the mounting bracket are aligned with the five locating pins on the side of the chassis.
2. Orient the slide-rail assembly so that the ball-bearing track is forward and locked in place.
3. Starting on either side of the rack, align the rear of the slide-rail assembly against the inside of the rear rack rail, and push until the assembly locks into place with an audible click.

Figure 2-1 Locking the Slide-Rail Assembly Against the Inside of the Rear Rack Rail



4. Align the front of the slide-rail assembly against the outside of the front rack rail, and push until the assembly locks into place and you hear the click.
5. Repeat steps 2 to 4 on the other side on the rack.

2.2.3.3 Installing the Server

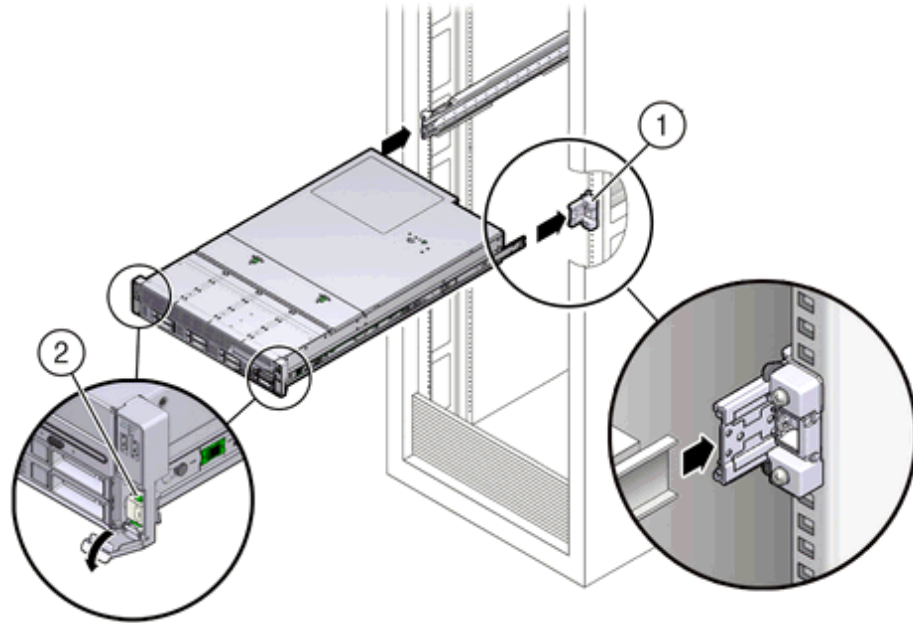
After preparing for the installation and installing the rack assembly, you then install the new server.

 **WARNING:**

- Installing a server requires a minimum of two people or a lift because of the weight of each server. Attempting this procedure alone can result in equipment damage, personal injury, or both.
- Always load equipment into the rack from the bottom up, so that the rack does not become top-heavy and tip over. Extend the rack anti-tip bar to prevent the rack from tipping during equipment installation.

1. Read the service label on the top cover of the server before installing a server into the rack.
2. Push the server into the slide rail assembly:
 - a. Push the slide rails into the slide rail assemblies as far as possible.
 - b. Position the server so the rear ends of the mounting brackets are aligned with the slide rail assemblies mounted in the equipment rack.

Figure 2-2 Aligning the Rear Ends of the Mounting Brackets with the Slide Rail Assemblies in the Rack



The callouts in the preceding image highlight the following:

- 1: Mounting bracket inserted into slide rail
- 2: Slide-rail release lever
- c. Insert the mounting brackets into the slide rails, and push the server into the rack until the mounting brackets encounter the slide rail stops, approximately 30 cm (12 inches).
- d. Simultaneously push down and hold the slide rail release levers on each mounting bracket while pushing the server into the rack.

 **Note:**

Oracle recommends that two people push the servers into the rack: one person to move the server in and out of the rack, and another person to watch the cables and cable management arm (CMA).

- e. Continue pushing until the slide rail locks on the front of the mounting brackets engage the slide rail assemblies, and you hear the click.
3. Cable the new server as described in [Cabling Exadata Storage Servers](#).

2.2.4 Cabling Database Servers

After the new database servers are installed, they need to be cabled to the existing equipment. The following procedure describes how to cable the new equipment in the rack. The images shown in the procedure are of a Sun Fire X4170 M2 Oracle Database Server.

 **Note:**

- The existing cable connections in the rack do not change.
- The blue cables connect to Oracle Database servers, and the black cables connect to Exadata Storage Servers. These network cables are for the NET0 Ethernet interface port.
- Attach and route the management cables on the CMA and rear panel one server at a time. Do not slide out more than one server at a time.
- Start from the bottom of the rack, and work upward. Route the cables through the CMA with the dongle on the top and power cables on the bottom.
- Longer hook and loop straps are needed when cabling three CAT5e cables or two TwinAx cables.

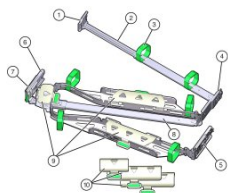
1. Connect the CAT5e cables, AC power cables, and USB to their respective ports on the rear of the server. Ensure the flat side of the dongle is flush against the CMA inner rail.

Figure 2-3 Cables at the Rear of the Server



2. Adjust the green cable management arm (CMA) brackets

Figure 2-4 Cable Management Arm (CMA) Brackets



Description of the CMA callouts in the preceding image"

- a. Connector A
- b. Front slide bar

- c. Velcro straps (6)
 - d. Connector B
 - e. Connector C
 - f. Connector D
 - g. Slide-rail latching bracket (used with connector D)
 - h. Rear slide bar
 - i. Cable covers
 - j. Cable covers
3. Attach the CMA to the server.
 4. Route the CAT5e and power cables through the wire clip.

Figure 2-5 Cables Routed Through the Cable Management Arm



5. Bend the CAT5e and power cables to enter the CMA, while adhering to the bend radius minimums.
6. Secure the CAT5e and power cables under the cable clasps.

Figure 2-6 Cables Secured under the Cable Clasps



7. Route the cables through the CMA, and secure them with hook and loop straps at equal intervals.

Figure 2-7 Cables Secured with Hook and Loop Straps at Regular Intervals



8. Connect the RDMA Network Fabric or TwinAx cables with the initial bend resting on the CMA. The TwinAx cables are for client access to the database servers.

Figure 2-8 RDMA Network Fabric or TwinAx Cables Positioned on the CMA



9. Secure the RDMA Network Fabric or TwinAx cables with hook and loop straps at equal intervals.

Figure 2-9 RDMA Network Fabric or TwinAx Cables Secured with Hook and Loop Straps at Regular Intervals



10. Route the fiber core cables.
11. Rest the cables over the green clasp on the CMA.
12. Attach the red ILOM cables to the database server.
13. Attach the network cables to the Oracle Database server.
14. Attach the cables from Oracle Database server to the RDMA Network Fabric switches.
15. Connect the orange Ethernet cable to the KVM switch.
16. Connect the red and blue Ethernet cables to the Cisco switch.
17. Verify operation of the slide rails and CMA for each server, as follows:

 **Note:**

Oracle recommends that two people do this step. One person to move the server in and out of the rack, and another person to observe the cables and CMA.

- a. Slowly pull the server out of the rack until the slide rails reach their stops.
 - b. Inspect the attached cables for any binding or kinks.
 - c. Verify the CMA extends fully from the slide rails.
18. Push the server back into the rack, as follows:
- a. Release the two sets of slide rail stops.
 - b. Push in both levers simultaneously, and slide the server into the rack. The first stop in the set are levers located on the inside of each slide rail, just behind the back panel of the server. The levers are labeled `PUSH`. The server slides approximately 46 cm (18 inches) and stop.
 - c. Verify the cables and CMA retract without binding.
 - d. Simultaneously push or pull both slide rail release buttons, and push the server completely into the rack until both slide rails engage. The second stop in the set are the slide rail release buttons located near the front of each mounting bracket.
19. Dress the cables, and then tie off the cables with the straps. Oracle recommends the cables should be dressed in bundles of eight or less.
20. Extend and then fully retract the server to check cable travel by sliding each server out and back fully to ensure that the cables are not binding or catching.
21. Repeat the procedure for the rest of the servers.
22. Connect the power cables to the power distribution units (PDUs). Ensure the breaker switches are in the OFF position before connecting the power cables. Do not plug the power cables into the facility receptacles at this time.

 **See Also:**

- *Oracle Exadata Database Machine System Overview* for cabling tables
- "[Reviewing the Cable Management Arm Guidelines](#)" for the bend radius minimums
- "[Floor and Underfloor Delivery of InfiniBand Cables](#)"

2.2.5 Cabling Storage Servers

After the new Storage Servers are installed, you need to connect them to the existing equipment.

The following procedure describes how to cable the new equipment in the rack.

 **Note:**

- The existing cable connections in the rack do not change.
- The blue cables connect to Oracle Database servers, and the black cables connect to Exadata Storage Servers. These network cables are for the NET0 Ethernet interface port.
- Attach and route the management cables on the CMA and rear panel one server at a time. Do not slide out more than one server at a time.
- Start from the bottom of the rack, and work upward.
- Longer hook and loop straps are needed when cabling three CAT5e cables or two TwinAx cables.

1. Attach a CMA to the server.
2. Insert the cables into their ports through the hook and loop straps, then route the cables into the CMA in this order:
 - a. Power
 - b. Ethernet
 - c. RDMA Network Fabric

Figure 2-10 Rear of the Server Showing Power and Network Cables



3. Route the cables through the CMA and secure them with hook and loop straps on both sides of each bend in the CMA.

Figure 2-11 Cables Routed Through the CMA and Secured with Hook and Loop Straps



4. Close the crossbar covers to secure the cables in the straightaway.
5. Verify operation of the slide rails and the CMA for each server:

 **Note:**

Oracle recommends that two people do this step: one person to move the server in and out of the rack, and another person to watch the cables and the CMA.

- a. Slowly pull the server out of the rack until the slide rails reach their stops.
- b. Inspect the attached cables for any binding or kinks.
- c. Verify that the CMA extends fully from the slide rails.
6. Push the server back into the rack:
 - a. Release the two sets of slide rail stops.
 - b. Locate the levers on the inside of each slide rail, just behind the back panel of the server. They are labeled PUSH.
 - c. Simultaneously push in both levers and slide the server into the rack, until it stops in approximately 46 cm (18 inches).
 - d. Verify that the cables and CMA retract without binding.
 - e. Locate the slide rail release buttons near the front of each mounting bracket.
 - f. Simultaneously push in both slide rail release buttons and slide the server completely into the rack, until both slide rails engage.
7. Dress the cables, and then tie off the cables with the straps. Oracle recommends that you dress the RDMA Network Fabric cables in bundles of eight or fewer.
8. Slide each server out and back fully to ensure that the cables are not binding or catching.
9. Repeat the procedure for all servers.
10. Connect the power cables to the power distribution units (PDUs). Ensure the breaker switches are in the OFF position before connecting the power cables. Do not plug the power cables into the facility receptacles now.

**See Also:**[Multi-Rack Cabling Tables](#)

Oracle Exadata Database Machine System Overview for the cabling tables for your system

2.2.6 Closing the Rack

After installing new equipment, you must replace the panels and close the rack.

There are two rack models in use with Oracle Exadata Racks. Refer to the appropriate documentation for the most up-to-date steps:

- *Oracle Rack Cabinet 1242 User's Guide* at http://docs.oracle.com/cd/E85660_01/html/E87280/index.html
- *Sun Rack II User's Guide* at <https://docs.oracle.com/cd/E19657-01/html/E29153/index.html>

The following steps provide an overview of the process.

1. Replace the rack front and rear doors as follows:
 - a. Retrieve the doors, and place them carefully on the door hinges.
 - b. Connect the front and rear door grounding strap to the frame.
 - c. Close the doors.
 - d. (Optional) Lock the doors. The keys are in the shipping kit.
2. (Optional) Replace the side panels, if they were removed for the upgrade, as follows:
 - a. Lift each side panel up and onto the side of the rack. The top of the rack should support the weight of the side panel. Ensure the panel fasteners line up with the grooves in the rack frame.
 - b. Turn each side panel fastener one-quarter turn clockwise using the side panel removal tool. Turn the fasteners next to the panel lock clockwise. There are 10 fasteners per side panel.
 - c. (Optional) Lock each side panel. The key is in the shipping kit. The locks are located on the bottom, center of the side panels.
 - d. Connect the grounding straps to the side panels.

After closing the rack, proceed to [Configuring the New Hardware](#) to configure the new hardware.

2.3 Extending a Rack by Adding Another Rack

You can extend your Oracle Exadata Rack by adding another rack and configuring the racks together.

- [Overview of Adding Another Rack to an Existing System](#)
Review the following notes before cabling racks together.

- [Cabling Two Racks Together](#)
Choose from the available methods based on your system specifications and operational requirements.
- [Cabling Several Racks Together](#)
By using the following procedures, you can add another rack to an existing multi-rack system.

2.3.1 Overview of Adding Another Rack to an Existing System

Review the following notes before cabling racks together.

- The procedures for extending racks with RoCE Network Fabric (X8M and later) are different than the procedures for racks with InfiniBand Network Fabric (X8 and earlier.)
- Racks with InfiniBand Network Fabric can be cabled together with no downtime. Depending on the procedure being used, racks with RoCE Network Fabric might require downtime when cabling racks together.
- Cabling within a live network must be done carefully in order to avoid potentially serious disruptions.
- There can be performance degradation while cabling the racks together. This degradation results from data retransmission due to packet loss and reduced network bandwidth when a cable is unplugged.
- Redundancy with the RDMA Network Fabric can be compromised while cabling the racks together. This occurs whenever the RDMA Network Fabric ports or switches are taken offline and all traffic must use the remaining switches.
- Only the existing racks are operational when adding racks. It is assumed that the servers on any new racks are initially powered down.
- The software running on the systems cannot have problems related to RDMA Network Fabric restarts. To verify the configuration, run `infinicheck` separately on each rack before connecting multiple racks together.
- It is assumed that each Oracle Exadata Rack has three RDMA Network Fabric switches already installed.
- The new racks have been configured with the appropriate IP addresses to be migrated into the expanded system prior to any cabling, and there are no duplicate IP addresses.
- Racks with RoCE Network Fabric use one loopback IP interface on each spine switch and two loopback IP interfaces on each leaf switch. The IP addressing scheme uses IANA 'Shared Address Space' 100.64.0.0/10. This ensures that there is no overlap with IPv4 addresses in the network using other schemes.
 - Leaf loopback0 IPs are assigned as 100.64.0.101, 100.64.0.102, 100.64.0.103, and so on.
 - Leaf loopback1 IPs are assigned as 100.64.1.101, 100.64.1.102, 100.64.1.103, and so on.
 - Spine loopback0 IPs are assigned as 100.64.0.201, 100.64.0.202, up to 100.64.0.208.

2.3.2 Cabling Two Racks Together

Choose from the available methods based on your system specifications and operational requirements.

- [Cabling Two RoCE Network Fabric Racks Together with No Down Time](#)
If your operational requirements cannot tolerate any scheduled down time, then choose from the following procedures to extend your existing RoCE Network Fabric rack by adding another rack.
- [Cabling Two RoCE Network Fabric Racks Together with Down Time](#)
If your operational requirements can tolerate some scheduled down time, then you can choose from the following procedures to extend your existing RoCE Network Fabric rack by adding another rack.
- [Cabling Two InfiniBand Network Fabric Racks Together](#)
Use this procedure to cable together two racks with InfiniBand Network Fabric.

2.3.2.1 Cabling Two RoCE Network Fabric Racks Together with No Down Time

If your operational requirements cannot tolerate any scheduled down time, then choose from the following procedures to extend your existing RoCE Network Fabric rack by adding another rack.

- [Extending an X9M Rack with No Down Time by Adding Another X9M Rack](#)
- [Extending an X8M Rack with No Down Time by Adding an X9M Rack](#)
- [Extending an X8M Rack with No Down Time by Adding Another X8M Rack](#)

2.3.2.1.1 Extending an X9M Rack with No Down Time by Adding Another X9M Rack

WARNING:

Take time to read and understand this procedure before implementation. Pay careful attention to the instructions that surround the command examples. A system outage may occur if the procedure is not applied correctly.

Note:

For additional background information, see [Understanding Multi-Rack Cabling for X9M Racks](#).

Use this procedure to extend a typical X9M rack by cabling it together with a second X9M rack. The primary rack (designated R1) and all of the systems it supports remain online throughout the procedure. At the beginning of the procedure, the additional rack (designated R2) is shut down.

The following is an outline of the procedure:

- Preparation (steps [1](#) and [2](#))

In this phase, you prepare the racks, switches, and cables. Also, you install and cable the spine switches in both racks.

- Configuration and Physical Cabling

In this phase, you reconfigure the leaf switches and finalize the cabling to the spine switches. These tasks are carefully orchestrated to avoid downtime on the primary system, as follows:

- Partially configure the lower leaf switches (step 3)

In this step, you reconfigure the switch ports on the lower leaf switches. There is no physical cabling performed in this step.

- Partially configure the upper leaf switches (step 4)

In this step, you reconfigure the switch ports on the upper leaf switches, remove the inter-switch cables that connect the leaf switches in both racks and connect the cables between the upper leaf switches and the spine switches.

- Finalize the lower leaf switches (step 5)

In this step, you finalize the switch port configuration on the lower leaf switches. You also complete the physical cabling by connecting the cables between the lower leaf switches and the spine switches.

- Finalize the upper leaf switches (step 6)

In this step, you finalize the switch port configuration on the upper leaf switches.

- Validation and Testing (steps 7 and 8)

In this phase, you validate and test the RoCE Network Fabric across both of the interconnect racks.

After completing the procedure, both racks share the RoCE Network Fabric, and the combined system is ready for further configuration. For example, you can extend existing disk groups and Oracle RAC databases to consume resources across both racks.

 **Note:**

- This procedure applies only to typical rack configurations that initially have leaf switches with the following specifications:
 - The inter-switch ports are ports 4 to 7, and ports 30 to 33.
 - The storage server ports are ports 8 to 14, and ports 23 to 29.
 - The database server ports are ports 15 to 22.

For other rack configurations (for example, X9M-8 systems with three database servers and 11 storage servers) a different procedure and different RoCE Network Fabric switch configuration files are required. Contact Oracle for further guidance.

- The procedure uses the following naming abbreviations and conventions:
 - The abbreviation for the existing rack is R1, and the new rack is R2.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies a spine switch.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the existing rack (R1).
- Most operations must be executed in multiple locations. For example, step [1.h](#) instructs you to update the firmware on all the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL). Pay attention to the instructions and keep track of your actions.

 **Tip:**

When a step must be performed on multiple switches, the instruction contains a list of the applicable switches. For example, (R1LL, R1UL, R2LL, and R2UL). You can use this list as a checklist to keep track of your actions.

- Perform operations sequentially, and complete every operation before proceeding. For example, run the entire command sequence at [3.a.i](#) as one operation and complete it before proceeding.
- All of commands that are run on a RoCE Network Fabric switch must be run while connected to the switch management interface as the switch administrator.

1. Prepare the systems.

- a. Position the new rack (R2) so that it is physically near the existing rack (R1).
The RDMA Network Fabric cables must be able to reach the switches in each rack.
For the required cross-rack cabling information, see [Two-Rack Cabling for X9M Racks](#).
- b. Power on all of the servers and network switches in the new rack (R2).

This includes the database servers, storage servers, RoCE Network Fabric leaf switches, and the Management Network Switch.

- c. Prepare the RoCE Network Fabric cables that you will use to interconnect the racks.

Label both ends of every cable.

For the required cross-rack cabling information, see [Two-Rack Cabling for X9M Racks](#).

- d. Connect the new rack (R2) to your existing management network.

Ensure that there are no IP address conflicts across the racks and that you can access the management interfaces on the RoCE Network Fabric switches.

- e. Ensure that you have a backup of the current switch configuration for each RoCE Network Fabric switch (R1LL, R1UL, R2LL, and R2UL).

See Backing Up Settings on the RoCE Network Fabric Switch in *Oracle Exadata Database Machine Maintenance Guide*.

- f. Download the required RoCE Network Fabric switch configuration files.

This procedure requires specific RoCE Network Fabric switch configuration files, which you must download from [My Oracle Support document 2704997.1](#).

 **WARNING:**

You must use different switch configuration files depending on whether your system uses Exadata Secure RDMA Fabric Isolation. Ensure that you download the correct archive that matches your system configuration.

For system configurations without Secure Fabric, download `online_multi-rack_14uplinks.zip`. For system configurations with Secure Fabric, download `online_SF_enabled_multi-rack_14uplinks.zip`.

Download and extract the archive containing the required RoCE Network Fabric switch configuration files. Place the files on a server with access to the management interfaces on the RoCE Network Fabric switches.

- g. Copy the required RoCE Network Fabric switch configuration files to the leaf switches on both racks.

You can use the following commands to copy the required configuration files to all of the RoCE Network Fabric switches on a system without Secure Fabric enabled:

- i.

```
# scp roce_multi_14uplinks_online_step3_R1_LL.cfg  
admin@R1LL_IP: /
```

- ii.

```
# scp roce_multi_14uplinks_online_step3_R2_LL.cfg  
admin@R2LL_IP: /
```

- iii. # scp roce_multi_14uplinks_online_step4_R1_UL.cfg admin@R1UL_IP:/
- iv. # scp roce_multi_14uplinks_online_step4_R2_UL.cfg admin@R2UL_IP:/
- v. # scp roce_multi_14uplinks_online_step5.cfg admin@R1LL_IP:/
- vi. # scp roce_multi_14uplinks_online_step5.cfg admin@R2LL_IP:/

On a system with Secure Fabric enabled, you can use the following commands:

- i. # scp roce_SF_multi_14uplinks_online_step3_R1_LL.cfg
admin@R1LL_IP:/
- ii. # scp roce_SF_multi_14uplinks_online_step3_R2_LL.cfg
admin@R2LL_IP:/
- iii. # scp roce_SF_multi_14uplinks_online_step4_R1_UL.cfg
admin@R1UL_IP:/
- iv. # scp roce_SF_multi_14uplinks_online_step4_R2_UL.cfg
admin@R2UL_IP:/
- v. # scp roce_SF_multi_14uplinks_online_step5.cfg admin@R1LL_IP:/
- vi. # scp roce_SF_multi_14uplinks_online_step5.cfg admin@R2LL_IP:/

In the above commands, substitute the appropriate IP address or host name where applicable. For example, in place of *R1LL_IP*, substitute the management IP address or host name for the lower leaf switch (LL) on the existing rack (R1).

 **Note:**

The command examples in the rest of this procedure use the configuration files for a system configuration without Secure Fabric enabled. If required, adjust the commands to use the Secure Fabric-enabled switch configuration files.

- h. Update the firmware to the latest available release on all of the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL).

See Updating RoCE Network Fabric Switch Firmware in *Oracle Exadata Database Machine Maintenance Guide*.

- i. Examine the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL) and confirm the port categories for the cabled ports.

Run the `show interface status` command on every RoCE Network Fabric leaf switch:

- i. R1LL# `show interface status`

ii. R1U1# **show interface status**

iii. R2L1# **show interface status**

iv. R2U1# **show interface status**

Examine the output and confirm the port categories as follows:

- Confirm that the inter-switch ports are ports 4 to 7, and ports 30 to 33.
- Confirm that the storage server ports are ports 8 to 14, and ports 23 to 29.
- Confirm that the database server ports are ports 15 to 22.

For example:

R1L1# **show interface status**

```
-----
-----
Port          Name          Status  Vlan    Duplex
Speed   Type
-----
-----
mgmt0         --            connected routed   full
1000         --
```

```
-----
-----
Port          Name          Status  Vlan    Duplex
Speed   Type
-----
-----
Eth1/1         --            xcvrAbsen 1      auto
auto         --
Eth1/2         --            xcvrAbsen 1      auto
auto         --
Eth1/3         --            xcvrAbsen 1      auto
auto         --
Eth1/4         ISL1          connected trunk   full
100G        QSFP-100G-CR4
Eth1/5         ISL2          connected trunk   full
100G        QSFP-100G-CR4
Eth1/6         ISL3          connected trunk   full
100G        QSFP-100G-CR4
Eth1/7         ISL4          connected trunk   full
100G        QSFP-100G-CR4
Eth1/8         celadm14     connected 3888   full
100G        QSFP-100G-CR4
Eth1/9         celadm13     connected 3888   full
100G        QSFP-100G-CR4
Eth1/10        celadm12     connected 3888   full
100G        QSFP-100G-CR4
Eth1/11        celadm11     connected 3888   full
100G        QSFP-100G-CR4
```

Eth1/12	celadm10	connected 3888	full	100G	
QSF100G-CR4					
Eth1/13	celadm09	connected 3888	full	100G	
QSF100G-CR4					
Eth1/14	celadm08	connected 3888	full	100G	
QSF100G-CR4					
Eth1/15	adm08	connected 3888	full	100G	
QSF100G-CR4					
Eth1/16	adm07	connected 3888	full	100G	
QSF100G-CR4					
Eth1/17	adm06	connected 3888	full	100G	
QSF100G-CR4					
Eth1/18	adm05	connected 3888	full	100G	
QSF100G-CR4					
Eth1/19	adm04	connected 3888	full	100G	
QSF100G-CR4					
Eth1/20	adm03	connected 3888	full	100G	
QSF100G-CR4					
Eth1/21	adm02	connected 3888	full	100G	
QSF100G-CR4					
Eth1/22	adm01	connected 3888	full	100G	
QSF100G-CR4					
Eth1/23	celadm07	connected 3888	full	100G	
QSF100G-CR4					
Eth1/24	celadm06	connected 3888	full	100G	
QSF100G-CR4					
Eth1/25	celadm05	connected 3888	full	100G	
QSF100G-CR4					
Eth1/26	celadm04	connected 3888	full	100G	
QSF100G-CR4					
Eth1/27	celadm03	connected 3888	full	100G	
QSF100G-CR4					
Eth1/28	celadm02	connected 3888	full	100G	
QSF100G-CR4					
Eth1/29	celadm01	connected 3888	full	100G	
QSF100G-CR4					
Eth1/30	ISL5	connected trunk	full	100G	
QSF100G-CR4					
Eth1/31	ISL6	connected trunk	full	100G	
QSF100G-CR4					
Eth1/32	ISL7	connected trunk	full	100G	
QSF100G-CR4					
Eth1/33	ISL8	connected trunk	full	100G	
QSF100G-CR4					
Eth1/34	--	xcvrAbsen 1	auto	auto	--
Eth1/35	--	xcvrAbsen 1	auto	auto	--
Eth1/36	--	xcvrAbsen 1	auto	auto	--
Po100	--	connected trunk	full	100G	--
Lo0	Routing loopback i	connected routed	auto	auto	--
Lo1	VTEP loopback inte	connected routed	auto	auto	--
Vlan1	--	down routed	auto	auto	--
nve1	--	connected --	auto	auto	--

- j. For each rack (R1 and R2), confirm the RoCE Network Fabric cabling by running the `verify_roce_cables.py` script.

The `verify_roce_cables.py` script uses two input files; one for database servers and storage servers (`nodes.rackN`), and another for switches (`switches.rackN`). In each file, every server or switch must be listed on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

See My Oracle Support document 2587717.1 for download and detailed usage instructions.

Run the `verify_roce_cables.py` script against both of the racks:

- i.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
```
- ii.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
```

Check that output in the `CABLE OK?` columns contains the `OK` status.

The following example shows the expected command results:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER)                LOWER LEAF (rack1sw-
rocea0) : CABLE OK?                UPPER LEAF (rack1sw-roceb0) :
CABLE OK?
-----
----- : -----
----- : -----
      Eth1/4 (ISL peer switch)   :          rack1sw-rocea0
Ethernet1/4 : OK                rack1sw-roceb0
Ethernet1/4 : OK
      Eth1/5 (ISL peer switch)   :          rack1sw-rocea0
Ethernet1/5 : OK                rack1sw-roceb0
Ethernet1/5 : OK
      Eth1/6 (ISL peer switch)   :          rack1sw-rocea0
Ethernet1/6 : OK                rack1sw-roceb0
Ethernet1/6 : OK
      Eth1/7 (ISL peer switch)   :          rack1sw-rocea0
Ethernet1/7 : OK                rack1sw-roceb0
Ethernet1/7 : OK
      Eth1/8 (RU39)              :          rack1celadm14
port-1 : OK                    rack1celadm14 port-2 : OK
      Eth1/9 (RU37)              :          rack1celadm13
port-1 : OK                    rack1celadm13 port-2 : OK
      Eth1/10 (RU35)             :          rack1celadm12
port-1 : OK                    rack1celadm12 port-2 : OK
      Eth1/11 (RU33)            :          rack1celadm11
port-1 : OK                    rack1celadm11 port-2 : OK
      Eth1/12 (RU31)            :          rack1celadm10
port-1 : OK                    rack1celadm10 port-2 : OK
      Eth1/13 (RU29)            :          rack1celadm09
port-1 : OK                    rack1celadm09 port-2 : OK
      Eth1/14 (RU27)            :          rack1celadm08
port-1 : OK                    rack1celadm08 port-2 : OK
      Eth1/15 (RU26)            :          rack1adm08
```

```

port-1 : OK
  Eth1/16 (RU25) : rack1adm8 port-2 : OK
port-1 : OK
  Eth1/17 (RU24) : rack1adm7 port-2 : OK
port-1 : OK
  Eth1/18 (RU23) : rack1adm6 port-2 : OK
port-1 : OK
  Eth1/19 (RU19) : rack1adm5 port-2 : OK
port-1 : OK
  Eth1/20 (RU18) : rack1adm4 port-2 : OK
port-1 : OK
  Eth1/21 (RU17) : rack1adm3 port-2 : OK
port-1 : OK
  Eth1/22 (RU16) : rack1adm2 port-2 : OK
port-1 : OK
  Eth1/23 (RU14) : rack1adm1 port-2 : OK
port-1 : OK
  Eth1/24 (RU12) : rack1celadm7 port-2 : OK
port-1 : OK
  Eth1/25 (RU10) : rack1celadm6 port-2 : OK
port-1 : OK
  Eth1/26 (RU08) : rack1celadm5 port-2 : OK
port-1 : OK
  Eth1/27 (RU06) : rack1celadm4 port-2 : OK
port-1 : OK
  Eth1/28 (RU04) : rack1celadm3 port-2 : OK
port-1 : OK
  Eth1/29 (RU02) : rack1celadm2 port-2 : OK
port-1 : OK
  Eth1/30 (ISL peer switch) : rack1celadm1 port-2 : OK
Ethernet1/30 : OK rack1sw-rocea0
  Eth1/31 (ISL peer switch) : rack1sw-roceb0 Ethernet1/30 : OK
Ethernet1/31 : OK rack1sw-rocea0
  Eth1/32 (ISL peer switch) : rack1sw-roceb0 Ethernet1/31 : OK
Ethernet1/32 : OK rack1sw-rocea0
  Eth1/33 (ISL peer switch) : rack1sw-roceb0 Ethernet1/32 : OK
Ethernet1/33 : OK rack1sw-roceb0 Ethernet1/33 : OK
  
```

- k. For each rack (R1 and R2), verify the RoCE Network Fabric operation by using the `infinicheck` command.
- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command.
 - Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric.
 - Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks.

In each command, the `hosts` input file (`hosts.rack1` and `hosts.rack2`) contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses (2 RoCE Network Fabric IP addresses for each database server), and the `cells` input file (`cells.rack1` and `cells.rack2`) contains a list of RoCE Network

Fabric IP addresses for the storage servers (2 RoCE Network Fabric IP addresses for each storage server).

- i. Use the following recommended command sequence on the existing rack (R1):

- i.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rack1 -c cells.rack1 -z
```
- ii.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rack1 -c cells.rack1 -s
```
- iii.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rack1 -c cells.rack1 -b
```

- ii. Use the following recommended command sequence on the new rack (R2):

- i.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rack2 -c cells.rack2 -z
```
- ii.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rack2 -c cells.rack2 -s
```
- iii.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rack2 -c cells.rack2 -b
```

The following example shows the expected command results for the final command in the sequence:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rackN -c cells.rackN -b

INFINICHECK
      [Network Connectivity, Configuration and
Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all
CELLs.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS].... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and
CELLs
[SUCCESS].... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS].... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs
and CELLS
[SUCCESS].... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value :
```

```

2300
[SUCCESS].... RoCE interface MTU settings look good
    Verifying switch advertised DSCP on all DBs and CELLS ports ( )
[SUCCESS].... Advertised DSCP settings from RoCE switch looks good
    ##### CONNECTIVITY TESTS #####
    [COMPUTE NODES -> STORAGE CELLS]
    (60 seconds approx.)
    (Will walk through QoS values: 0-6) [SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
    [COMPUTE NODES -> COMPUTE NODES]
...
  
```

2. Install the spine switches (R1SS and R2SS).

- a. Physically install and power up the spine switches in the existing rack (R1SS) and the new rack (R2SS).
 - i. Physically install each spine switch in RU1.
 - ii. For each spine switch, ensure that the management Ethernet interface is connected to the management network and then supply power.
 - iii. On each spine switch, perform the initial configuration steps outlined in *Configuring the Cisco Nexus C9336C-FX2 Switch*. Skip the step for applying the golden configuration settings as you will do this later.
 - iv. For each spine switch, perform a ping test to the management Ethernet interface to ensure that the switch is online and accessible.

b. Apply the golden configuration settings to the new spine switches.

See *Applying Golden Configuration Settings on RoCE Network Fabric Switches in Oracle Exadata Database Machine Maintenance Guide*.

You can use the instance of `patchmgr` that you previously used to update the firmware on the leaf switches (in step 1.h).

Use a switch list file (`spines.lst`) to apply the golden configuration settings to both spine switches using one `patchmgr` command:

```

# cat spines.lst
R1SS_IP:mspine.201
R2SS_IP:mspine.202

# ./patchmgr --roceswitches spines.lst --apply-config -log_dir /tmp/
spinelogs
  
```

 **Note:**

In the switch list file, `R1SS_IP` is the management IP address or host name for the spine switch on the existing rack (R1SS) and `R2SS_IP` is the management IP address or host name for the spine switch on the new rack (R2SS).

c. Upgrade the firmware on the spine switches.

See *Updating RoCE Network Fabric Switch Firmware in Oracle Exadata Database Machine Maintenance Guide*.

You can use the instance of `patchmgr` that you used in the previous step.

Use a switch list file (`spines.lst`) to perform the firmware upgrade on both spine switches using one `patchmgr` command:

```
# cat spines.lst
R1SS_IP:mspine.201
R2SS_IP:mspine.202

# ./patchmgr --roceswitches spines.lst --upgrade -log_dir /tmp/
spinelogs
```

 **Note:**

In the switch list file, `R1SS_IP` is the management IP address or host name for the spine switch on the existing rack (R1SS) and `R2SS_IP` is the management IP address or host name for the spine switch on the new rack (R2SS).

- d. Connect the RoCE Network Fabric cables to the spine switches (R1SS and R2SS).

 **WARNING:**

**At this stage, only connect the cables to the spine switches.
DO NOT CONNECT ANY OF THE CABLES TO THE LEAF
SWITCHES.**

Use the cabled that you prepared earlier (in step 1.c).

For the required cross-rack cabling information, see [Two-Rack Cabling for X9M Racks](#).

3. Perform the first round of configuration on the lower leaf switches (R1LL and R2LL).

Perform this step on the lower leaf switches (R1LL and R2LL) only.

 **Note:**

During this step, the lower leaf switch ports are shut down. While the R1LL ports are down, R1UL exclusively supports the RoCE Network Fabric. During this time, there is no redundancy in the RoCE Network Fabric, and availability cannot be maintained if R1UL goes down.

- a. Shut down the switch ports on the lower leaf switches (R1LL and R2LL).

i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/1-36
R1LL(config-if-range)# shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/1-36
R2LL(config-if-range)# shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

b. Reconfigure the lower leaf switch ports (R1LL and R2LL) .

For each switch, you must use the correct corresponding switch configuration file, which you earlier copied to the switch (in step 1.g).

i. On R1LL, the switch configuration file name must end with `step3_R1_LL.cfg`:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step3_R1_LL.cfg | grep 'none'
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. On R2LL, the switch configuration file name must end with `step3_R2_LL.cfg`:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step3_R2_LL.cfg | grep 'none'
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
```

```
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- c. Start the inter-switch ports on the lower leaf switches (R1LL and R2LL) .
 - i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/1-7, ethernet 1/30-36
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/1-7, ethernet 1/30-36
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- d. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- e. Verify the status of the inter-switch ports on the lower leaf switches (R1LL and R2LL) .

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**
- ii. R2LL# **show interface status**

Examine the output to ensure that the inter-switch ports are connected.

For example:

R1LL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
mgmt0         --           connected  routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
Eth1/1        --           xcvrAbsen  1         auto    auto    --
Eth1/2        --           xcvrAbsen  1         auto    auto    --
Eth1/3        --           xcvrAbsen  1         auto    auto    --
Eth1/4        ISL1         connected  trunk     full    100G
QSF1P-100G-CR4
Eth1/5        ISL2         connected  trunk     full    100G
QSF1P-100G-CR4
Eth1/6        ISL3         connected  trunk     full    100G
QSF1P-100G-CR4
Eth1/7        ISL4         connected  trunk     full    100G
QSF1P-100G-CR4
Eth1/8        celadm14    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/9        celadm13    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/10       celadm12    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/11       celadm11    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/12       celadm10    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/13       celadm09    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/14       celadm08    disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/15       adm08       disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/16       adm07       disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/17       adm06       disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/18       adm05       disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/19       adm04       disabled   3888     full    100G
QSF1P-100G-CR4
Eth1/20       adm03       disabled   3888     full    100G
QSF1P-100G-CR4

```

```

Eth1/21      adm02          disabled 3888    full
100G QSFP-100G-CR4
Eth1/22      adm01          disabled 3888    full
100G QSFP-100G-CR4
Eth1/23      celadm07       disabled 3888    full
100G QSFP-100G-CR4
Eth1/24      celadm06       disabled 3888    full
100G QSFP-100G-CR4
Eth1/25      celadm05       disabled 3888    full
100G QSFP-100G-CR4
Eth1/26      celadm04       disabled 3888    full
100G QSFP-100G-CR4
Eth1/27      celadm03       disabled 3888    full
100G QSFP-100G-CR4
Eth1/28      celadm02       disabled 3888    full
100G QSFP-100G-CR4
Eth1/29      celadm01       disabled 3888    full
100G QSFP-100G-CR4
Eth1/30      ISL5           connected trunk full
100G QSFP-100G-CR4
Eth1/31      ISL6           connected trunk full
100G QSFP-100G-CR4
Eth1/32      ISL7           connected trunk full
100G QSFP-100G-CR4
Eth1/33      ISL8           connected trunk full
100G QSFP-100G-CR4
Eth1/34      --            xcvrAbsen 1    auto
auto --
Eth1/35      --            xcvrAbsen 1    auto
auto --
Eth1/36      --            xcvrAbsen 1    auto
auto --
Po100       --            connected trunk full
100G --
Lo0         Routing loopback i connected routed auto
auto --
Lo1         VTEP loopback inte connected routed auto
auto --
Vlan1      --            down          routed auto
auto --
nve1       --            connected --   auto
auto --
  
```

f. Start the storage server ports on the lower leaf switches (R1LL and R2LL) .

i. On R1LL:

```

R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
  
```

```
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- g. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- h. Verify the status of the storage server ports on the lower leaf switches (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**
- ii. R2LL# **show interface status**

Examine the output to ensure that the storage server ports are connected.

For example:

```
R1LL# show interface status
```

```
-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
mgmt0         --            connected   routed    full   1000   --
-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
Eth1/1        --            xcvrAbsen  1         auto   auto   --
Eth1/2        --            xcvrAbsen  1         auto   auto   --
Eth1/3        --            xcvrAbsen  1         auto   auto   --
Eth1/4        ISL1          connected   trunk     full   100G
QSFP-100G-CR4
Eth1/5        ISL2          connected   trunk     full   100G
```



```

QSF-100G-CR4
Eth1/6      ISL3      connected trunk    full
100G      QSF-100G-CR4
Eth1/7      ISL4      connected trunk    full
100G      QSF-100G-CR4
Eth1/8      celadm14  connected 3888     full
100G      QSF-100G-CR4
Eth1/9      celadm13  connected 3888     full
100G      QSF-100G-CR4
Eth1/10     celadm12  connected 3888     full
100G      QSF-100G-CR4
Eth1/11     celadm11  connected 3888     full
100G      QSF-100G-CR4
Eth1/12     celadm10  connected 3888     full
100G      QSF-100G-CR4
Eth1/13     celadm09  connected 3888     full
100G      QSF-100G-CR4
Eth1/14     celadm08  connected 3888     full
100G      QSF-100G-CR4
Eth1/15     adm08     disabled  3888     full
100G      QSF-100G-CR4
Eth1/16     adm07     disabled  3888     full
100G      QSF-100G-CR4
Eth1/17     adm06     disabled  3888     full
100G      QSF-100G-CR4
Eth1/18     adm05     disabled  3888     full
100G      QSF-100G-CR4
Eth1/19     adm04     disabled  3888     full
100G      QSF-100G-CR4
Eth1/20     adm03     disabled  3888     full
100G      QSF-100G-CR4
Eth1/21     adm02     disabled  3888     full
100G      QSF-100G-CR4
Eth1/22     adm01     disabled  3888     full
100G      QSF-100G-CR4
Eth1/23     celadm07  connected 3888     full
100G      QSF-100G-CR4
Eth1/24     celadm06  connected 3888     full
100G      QSF-100G-CR4
Eth1/25     celadm05  connected 3888     full
100G      QSF-100G-CR4
Eth1/26     celadm04  connected 3888     full
100G      QSF-100G-CR4
Eth1/27     celadm03  connected 3888     full
100G      QSF-100G-CR4
Eth1/28     celadm02  connected 3888     full
100G      QSF-100G-CR4
Eth1/29     celadm01  connected 3888     full
100G      QSF-100G-CR4
Eth1/30     ISL5      connected trunk    full
100G      QSF-100G-CR4
Eth1/31     ISL6      connected trunk    full
100G      QSF-100G-CR4
Eth1/32     ISL7      connected trunk    full
100G      QSF-100G-CR4
  
```

Eth1/33	ISL8	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/34	--	xcvrAbsen 1	auto	auto	--
Eth1/35	--	xcvrAbsen 1	auto	auto	--
Eth1/36	--	xcvrAbsen 1	auto	auto	--
Po100	--	connected trunk	full	100G	--
Lo0	Routing loopback i	connected routed	auto	auto	--
Lo1	VTEP loopback inte	connected routed	auto	auto	--
Vlan1	--	down routed	auto	auto	--
nve1	--	connected --	auto	auto	--

- i. Start the database server ports on the lower leaf switches (R1LL and R2LL).

- i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/15-22
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/15-22
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- j. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- k. Verify the status of the database server ports on the lower leaf switches (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**

- ii. R2LL# **show interface status**

Examine the output to ensure that the database server ports are connected.

For example:

R1LL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex
Speed      Type
-----
mgmt0         --           connected   routed    full
1000         --
  
```

```

-----
-----
Port          Name          Status      Vlan      Duplex
Speed      Type
-----
Eth1/1         --           xcvrAbsen  1         auto
auto         --
Eth1/2         --           xcvrAbsen  1         auto
auto         --
Eth1/3         --           xcvrAbsen  1         auto
auto         --
Eth1/4         ISL1         connected   trunk     full
100G         QSFP-100G-CR4
Eth1/5         ISL2         connected   trunk     full
100G         QSFP-100G-CR4
Eth1/6         ISL3         connected   trunk     full
100G         QSFP-100G-CR4
Eth1/7         ISL4         connected   trunk     full
100G         QSFP-100G-CR4
Eth1/8         celadm14     connected   3888     full
100G         QSFP-100G-CR4
Eth1/9         celadm13     connected   3888     full
100G         QSFP-100G-CR4
Eth1/10        celadm12     connected   3888     full
100G         QSFP-100G-CR4
Eth1/11        celadm11     connected   3888     full
100G         QSFP-100G-CR4
Eth1/12        celadm10     connected   3888     full
100G         QSFP-100G-CR4
Eth1/13        celadm09     connected   3888     full
100G         QSFP-100G-CR4
Eth1/14        celadm08     connected   3888     full
100G         QSFP-100G-CR4
Eth1/15        adm08        connected   3888     full
100G         QSFP-100G-CR4
Eth1/16        adm07        connected   3888     full
100G         QSFP-100G-CR4
Eth1/17        adm06        connected   3888     full
100G         QSFP-100G-CR4
Eth1/18        adm05        connected   3888     full
100G         QSFP-100G-CR4
  
```

Eth1/19	adm04	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/20	adm03	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/21	adm02	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/22	adm01	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/23	celadm07	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/24	celadm06	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/25	celadm05	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/26	celadm04	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/27	celadm03	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/28	celadm02	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/29	celadm01	connected 3888	full	100G	
QSFP-100G-CR4					
Eth1/30	ISL5	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/31	ISL6	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/32	ISL7	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/33	ISL8	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/34	--	xcvrAbsen 1	auto	auto	--
Eth1/35	--	xcvrAbsen 1	auto	auto	--
Eth1/36	--	xcvrAbsen 1	auto	auto	--
Po100	--	connected trunk	full	100G	--
Lo0	Routing loopback i	connected routed	auto	auto	--
Lo1	VTEP loopback inte	connected routed	auto	auto	--
Vlan1	--	down routed	auto	auto	--
nve1	--	connected --	auto	auto	--

 **Note:**

Before proceeding, ensure that you have completed all of the actions in step 3 on both lower leaf switches (R1LL and R2LL). If not, then ensure that you go back and perform the missing actions.

4. Perform the first round of configuration on the upper leaf switches (R1UL and R2UL). Perform this step on the upper leaf switches (R1UL and R2UL) only.

 **Note:**

At the start of this step, the upper leaf switch ports are shut down. While the R1UL ports are down, R1LL exclusively supports the RoCE Network Fabric on the existing rack. During this time, there is no redundancy in the RoCE Network Fabric, and availability cannot be maintained if R1LL goes down.

a. Shut down the upper leaf switch ports (R1UL and R2UL).

i. On R1UL:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/1-36
R1UL(config-if-range)# shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

ii. Repeat the command sequence on R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/1-36
R2UL(config-if-range)# shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

b. On both racks, remove the inter-switch links between the leaf switches (R1LL to R1UL, and R2LL to R2UL).

On every leaf switch, remove the cables for the inter-switch links:

- i. On R1LL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
- ii. On R1UL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
- iii. On R2LL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
- iv. On R2UL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.

- c. On both racks, cable the upper leaf switch to both of the spine switches (R1UL and R2UL to R1SS and R2SS).

Connect the cables from the spine switches that you prepared earlier (in step 2.d).

Cable the switches as described in [Two-Rack Cabling for X9M Racks](#):

- i. On R1UL, cable ports 01, 02, 03, 04, 05, 06, 07, 30, 31, 32, 33, 34, 35, and 36 to R1SS and R2SS.
 - ii. On R2UL, cable ports 01, 02, 03, 04, 05, 06, 07, 30, 31, 32, 33, 34, 35, and 36 to R1SS and R2SS.
- d. Reconfigure the upper leaf switch ports (R1UL and R2UL).

For each switch, you must use the correct corresponding switch configuration file, which you earlier copied to the switch (in step 1.g):

- i. On R1UL, the switch configuration file name must end with `step4_R1_UL.cfg`:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step4_R1_UL.cfg | grep 'none'
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. On R2UL, the switch configuration file name must end with `step4_R2_UL.cfg`:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step4_R2_UL.cfg | grep 'none'
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- e. Check the status of the RoCE Network Fabric ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

- i. R1UL# **show interface status**

ii. R2UL# **show interface status**

Examine the output to ensure that all of the cabled ports are disabled.

For example:

R1UL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex
Speed        Type
-----
mgmt0         --            connected   routed    full
1000         --
-----
-----
Port          Name          Status      Vlan      Duplex
Speed        Type
-----
Eth1/1       RouterPort1   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/2       RouterPort2   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/3       RouterPort3   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/4       RouterPort4   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/5       RouterPort5   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/6       RouterPort6   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/7       RouterPort7   disabled    routed    full
100G        QSFP-100G-CR4
Eth1/8       celadm14      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/9       celadm13      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/10      celadm12      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/11      celadm11      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/12      celadm10      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/13      celadm09      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/14      celadm08      disabled    3888     full
100G        QSFP-100G-CR4
Eth1/15      adm08         disabled    3888     full
100G        QSFP-100G-CR4
Eth1/16      adm07         disabled    3888     full
100G        QSFP-100G-CR4
-----
  
```

```

Eth1/17      adm06      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/18      adm05      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/19      adm04      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/20      adm03      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/21      adm02      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/22      adm01      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/23      celadm07   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/24      celadm06   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/25      celadm05   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/26      celadm04   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/27      celadm03   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/28      celadm02   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/29      celadm01   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/30      RouterPort8 disabled routed   full 100G
QSFP-100G-CR4
Eth1/31      RouterPort9 disabled routed   full 100G
QSFP-100G-CR4
Eth1/32      RouterPort10 disabled routed   full 100G
QSFP-100G-CR4
Eth1/33      RouterPort11 disabled routed   full 100G
QSFP-100G-CR4
Eth1/34      RouterPort12 disabled routed   full 100G
QSFP-100G-CR4
Eth1/35      RouterPort13 disabled routed   full 100G
QSFP-100G-CR4
Eth1/36      RouterPort14 disabled routed   full 100G
QSFP-100G-CR4
Lo0          Routing loopback i connected routed   auto auto  --
Lo1          VTEP loopback inte connected routed   auto auto  --
Vlan1       --          down    routed   auto auto  --
nve1        --          connected --      auto auto  --

```

 **Note:**

Before proceeding, ensure that you have completed all of the actions to this point in step 4 on both upper leaf switches (R1UL and R2UL). If not, then ensure that you go back and perform the missing actions.

- f. Verify the configuration of the upper leaf switches.

You can use the instance of `patchmgr` that you previously used to update the switch firmware (in step 1.h).

Use a switch list file (`ul.lst`) to check both upper leaf switches using one `patchmgr` command:

```
# cat ul.lst
R1UL_IP:mleaf_u14.102
R2UL_IP:mleaf_u14.104
```

On a system with Secure Fabric enabled, use the `msfleaf_u14` tag in the switch list file:

```
# cat ul.lst
R1UL_IP:msfleaf_u14.102
R2UL_IP:msfleaf_u14.104
```

The following shows the recommended command and an example of the expected results:

```
# ./patchmgr --roceswitches ul.lst --verify-config -
log_dir /tmp/log
2020-08-10 13:40:09 -0700      :Working: Initiating config
verification... Expect up to 6 minutes for each switch
Mon Aug 10 13:40:13 PDT 2020 1 of 4 :Verifying config on
switch ...
...
Mon Aug 10 13:40:32 PDT 2020:      [INFO      ] Config matches
template: ...
Mon Aug 10 13:40:32 PDT 2020:      [SUCCESS   ] Config validation
successful!
2020-08-10 13:40:32 -0700
  Config check on RoCE switch(es)
2020-08-10 13:40:32 -0700
  Completed run of command: ./patchmgr --roceswitches ul.lst --
verify-config -log_dir /tmp/log
2020-08-10 13:40:32 -0700      :INFO      : config attempted on
nodes in file ul.lst: [R1UL_IP R2UL_IP]
2020-08-10 13:40:32 -0700      :INFO      : For details, check
the following files in /tmp/log:
2020-08-10 13:40:32 -0700      :INFO      : -
updateRoceSwitch.log
2020-08-10 13:40:32 -0700      :INFO      : -
updateRoceSwitch.trc
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.stdout
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.stderr
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.log
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.trc
2020-08-10 13:40:32 -0700      :INFO      : Exit status:0
2020-08-10 13:40:32 -0700      :INFO      : Exiting.
```

In the command output, verify that the switch configuration is good for both upper leaf switches. You can ignore messages about the ports that are down.

5. Finalize the configuration of the lower leaf switches (R1LL and R2LL).

Perform this step on the lower leaf switches (R1LL and R2LL) only.

a. Reconfigure the lower leaf switch ports (R1LL and R2LL).

Run the following command sequence on both of the lower leaf switches (R1LL and R2LL).

You must use the correct switch configuration file, which you earlier copied to the switch (in step 1.g). In this step, the configuration file name must end with `step5.cfg`.

i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step5.cfg | grep 'none'
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step5.cfg | grep 'none'
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

b. On both racks, cable the lower leaf switch to both of the spine switches (R1LL and R2LL to R1SS and R2SS).

Connect the cables from the spine switches that you prepared earlier (in step 2.d).

Cable the switches as described in [Two-Rack Cabling for X9M Racks](#):

- i. On R1LL, cable ports 01, 02, 03, 04, 05, 06, 07, 30, 31, 32, 33, 34, 35, and 36 to R1SS and R2SS.
 - ii. On R2LL, cable ports 01, 02, 03, 04, 05, 06, 07, 30, 31, 32, 33, 34, 35, and 36 to R1SS and R2SS.
- c. On the lower leaf switches, verify that all of the cabled RoCE Network Fabric ports are connected (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

i. R1LL# `show interface status`

ii. R2LL# `show interface status`

Examine the output to ensure that all of the cabled ports are connected.

For example:

R1LL# `show interface status`

```

-----
-----
Port          Name          Status  Vlan    Duplex
Speed  Type
-----
-----
mgmt0        --          connected routed    full
1000        --
-----
-----
Port          Name          Status  Vlan    Duplex
Speed  Type
-----
-----
Eth1/1        RouterPort1   connected routed    full
100G  QSFP-100G-CR4
Eth1/2        RouterPort2   connected routed    full
100G  QSFP-100G-CR4
Eth1/3        RouterPort3   connected routed    full
100G  QSFP-100G-CR4
Eth1/4        RouterPort4   connected routed    full
100G  QSFP-100G-CR4
Eth1/5        RouterPort5   connected routed    full
100G  QSFP-100G-CR4
Eth1/6        RouterPort6   connected routed    full
100G  QSFP-100G-CR4
Eth1/7        RouterPort7   connected routed    full
100G  QSFP-100G-CR4
Eth1/8        celadm14      connected 3888    full
100G  QSFP-100G-CR4
Eth1/9        celadm13      connected 3888    full
100G  QSFP-100G-CR4
Eth1/10       celadm12      connected 3888    full
100G  QSFP-100G-CR4
Eth1/11       celadm11      connected 3888    full
100G  QSFP-100G-CR4
Eth1/12       celadm10      connected 3888    full
100G  QSFP-100G-CR4
Eth1/13       celadm09      connected 3888    full
100G  QSFP-100G-CR4
Eth1/14       celadm08      connected 3888    full
100G  QSFP-100G-CR4

```

```

Eth1/15      adm08      connected 3888      full 100G
QSFP-100G-CR4
Eth1/16      adm07      connected 3888      full 100G
QSFP-100G-CR4
Eth1/17      adm06      connected 3888      full 100G
QSFP-100G-CR4
Eth1/18      adm05      connected 3888      full 100G
QSFP-100G-CR4
Eth1/19      adm04      connected 3888      full 100G
QSFP-100G-CR4
Eth1/20      adm03      connected 3888      full 100G
QSFP-100G-CR4
Eth1/21      adm02      connected 3888      full 100G
QSFP-100G-CR4
Eth1/22      adm01      connected 3888      full 100G
QSFP-100G-CR4
Eth1/23      celadm07   connected 3888      full 100G
QSFP-100G-CR4
Eth1/24      celadm06   connected 3888      full 100G
QSFP-100G-CR4
Eth1/25      celadm05   connected 3888      full 100G
QSFP-100G-CR4
Eth1/26      celadm04   connected 3888      full 100G
QSFP-100G-CR4
Eth1/27      celadm03   connected 3888      full 100G
QSFP-100G-CR4
Eth1/28      celadm02   connected 3888      full 100G
QSFP-100G-CR4
Eth1/29      celadm01   connected 3888      full 100G
QSFP-100G-CR4
Eth1/30      RouterPort8  connected routed      full 100G
QSFP-100G-CR4
Eth1/31      RouterPort9  connected routed      full 100G
QSFP-100G-CR4
Eth1/32      RouterPort10 connected routed      full 100G
QSFP-100G-CR4
Eth1/33      RouterPort11 connected routed      full 100G
QSFP-100G-CR4
Eth1/34      RouterPort12 connected routed      full 100G
QSFP-100G-CR4
Eth1/35      RouterPort13 connected routed      full 100G
QSFP-100G-CR4
Eth1/36      RouterPort14 connected routed      full 100G
QSFP-100G-CR4
Lo0          Routing loopback i connected routed      auto auto      --
Lo1          VTEP loopback inte connected routed      auto auto      --
Vlan1       --          down      routed      auto auto      --
nve1        --          connected --          auto auto      --
  
```

 **Note:**

Before proceeding, ensure that you have completed all of the actions to this point in step 5 on both lower leaf switches (R1LL and R2LL). If not, then ensure that you go back and perform the missing actions.

d. Verify the configuration of the lower leaf switches.

You can use the instance of `patchmgr` that you previously used to update the switch firmware (in step 1.h).

Use a switch list file (`ll.lst`) to check both lower leaf switches using one `patchmgr` command:

```
# cat ll.lst
R1LL_IP:mleaf_u14.101
R2LL_IP:mleaf_u14.103
```

On a system with Secure Fabric enabled, use the `msfleaf_u14` tag in the switch list file:

```
# cat ll.lst
R1LL_IP:msfleaf_u14.101
R2LL_IP:msfleaf_u14.103
```

The following shows the recommended command and an example of the expected results:

```
# ./patchmgr --roceswitches ll.lst --verify-config -
log_dir /tmp/log
2020-08-10 13:45:09 -0700      :Working: Initiating config
verification... Expect up to 6 minutes for each switch
Mon Aug 10 13:45:13 PDT 2020 1 of 4 :Verifying config on
switch ...
...
Mon Aug 10 13:45:32 PDT 2020:      [INFO      ] Config matches
template: ...
Mon Aug 10 13:45:32 PDT 2020:      [SUCCESS   ] Config validation
successful!
2020-08-10 13:45:32 -0700
  Config check on RoCE switch(es)
2020-08-10 13:45:32 -0700
  Completed run of command: ./patchmgr --roceswitches ll.lst --
verify-config -log_dir /tmp/log
2020-08-10 13:45:32 -0700      :INFO      : config attempted on
nodes in file ll.lst: [R1LL_IP R2LL_IP]
2020-08-10 13:45:32 -0700      :INFO      : For details, check
the following files in /tmp/log:
2020-08-10 13:45:32 -0700      :INFO      : -
updateRoceSwitch.log
2020-08-10 13:45:32 -0700      :INFO      : -
updateRoceSwitch.trc
2020-08-10 13:45:32 -0700      :INFO      : - patchmgr.stdout
```

```

2020-08-10 13:45:32 -0700      :INFO  : - patchmgr.stderr
2020-08-10 13:45:32 -0700      :INFO  : - patchmgr.log
2020-08-10 13:45:32 -0700      :INFO  : - patchmgr.trc
2020-08-10 13:45:32 -0700      :INFO  : Exit status:0
2020-08-10 13:45:32 -0700      :INFO  : Exiting.

```

In the command output, verify that the switch configuration is good for both lower leaf switches.

- e. Verify that nve is up on the lower leaf switches (R1LL and R2LL).

Run the following command on each lower leaf switch and examine the output:

- i. R1LL# **show nve peers**

- ii. R2LL# **show nve peers**

At this point, you should see one nve peer with State=Up.

For example:

```

R1LL# show nve peers
Interface      Peer-IP State LearnType  Uptime      Router-Mac
-----
nve1           100.64.1.103  Up        CP 00:04:29      n/a

```

- f. Verify that BGP is up on the lower leaf switches (R1LL and R2LL).

Run the following command on each lower leaf switch and examine the output:

- i. R1LL# **show logging log | grep BGP**

- ii. R2LL# **show logging log | grep BGP**

Look for two entries with Up in the rightmost column that are associated with different IP addresses.

For example:

```

R1LL# show logging log | grep BGP
2020 Aug 10 13:47:13 R1LL %BGP-5-ADJCHANGE: bgp- [29342] (default)
neighbor 100.64.0.201 Up
2020 Aug 10 13:47:24 R1LL %BGP-5-ADJCHANGE: bgp- [29342] (default)
neighbor 100.64.0.202 Up

```

6. Finalize the configuration of the upper leaf switches (R1UL and R2UL).

Perform this step on the upper leaf switches (R1UL and R2UL) only.

- a. Start the inter-switch ports on the upper leaf switches (R1UL and R2UL).

- i. On R1UL:

```

R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/1-7, ethernet 1/30-36
R1UL(config-if-range)# no shut

```

```
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. Repeat the command sequence on R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/1-7, ethernet 1/30-36
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

- b. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- c. Verify the status of the inter-switch ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

- i. R1UL# `show interface status`
- ii. R2UL# `show interface status`

Examine the output to ensure that the inter-switch ports are connected.

For example:

```
R1UL# show interface status
```

Port	Name	Status	Vlan	Duplex
mgmt0	--	connected	routed	full
1000	--			

Port	Name	Status	Vlan	Duplex
Eth1/1	RouterPort1	connected	routed	full

100G	QSFP-100G-CR4				
Eth1/2	RouterPort2	connected	routed	full	100G
QSFP-100G-CR4					
Eth1/3	RouterPort3	connected	routed	full	100G
QSFP-100G-CR4					
Eth1/4	RouterPort4	connected	routed	full	100G
QSFP-100G-CR4					
Eth1/5	RouterPort5	connected	routed	full	100G
QSFP-100G-CR4					
Eth1/6	RouterPort6	connected	routed	full	100G
QSFP-100G-CR4					
Eth1/7	RouterPort7	connected	routed	full	100G
QSFP-100G-CR4					
Eth1/8	celadm14	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/9	celadm13	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/10	celadm12	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/11	celadm11	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/12	celadm10	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/13	celadm09	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/14	celadm08	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/15	adm08	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/16	adm07	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/17	adm06	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/18	adm05	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/19	adm04	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/20	adm03	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/21	adm02	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/22	adm01	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/23	celadm07	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/24	celadm06	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/25	celadm05	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/26	celadm04	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/27	celadm03	disabled	3888	full	100G
QSFP-100G-CR4					
Eth1/28	celadm02	disabled	3888	full	100G
QSFP-100G-CR4					


```

Eth1/29      celadm01      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/30      RouterPort8      connected routed      full
100G      QSFP-100G-CR4
Eth1/31      RouterPort9      connected routed      full
100G      QSFP-100G-CR4
Eth1/32      RouterPort10     connected routed      full
100G      QSFP-100G-CR4
Eth1/33      RouterPort11     connected routed      full
100G      QSFP-100G-CR4
Eth1/34      RouterPort12     connected routed      full
100G      QSFP-100G-CR4
Eth1/35      RouterPort13     connected routed      full
100G      QSFP-100G-CR4
Eth1/36      RouterPort14     connected routed      full
100G      QSFP-100G-CR4
Lo0         Routing loopback i connected routed      auto
auto      --
Lo1         VTEP loopback inte connected routed      auto
auto      --
Vlan1       --                down      routed      auto
auto      --
nve1       --                connected --        auto
auto      --
  
```

d. Start the storage server ports on the upper leaf switches (R1UL and R2UL).

i. On R1UL:

```

R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
  
```

ii. Repeat the command sequence on R2UL:

```

R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
  
```

- e. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- f. Verify the status of the storage server ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

i. R1UL# `show interface status`

ii. R2UL# `show interface status`

Examine the output to ensure that the storage server ports are `connected`.

For example:

R1UL# `show interface status`

```

-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
mgmt0         --            connected   routed    full 1000  --
-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
Eth1/1        RouterPort1   connected   routed    full 100G
QSFP-100G-CR4
Eth1/2        RouterPort2   connected   routed    full 100G
QSFP-100G-CR4
Eth1/3        RouterPort3   connected   routed    full 100G
QSFP-100G-CR4
Eth1/4        RouterPort4   connected   routed    full 100G
QSFP-100G-CR4
Eth1/5        RouterPort5   connected   routed    full 100G
QSFP-100G-CR4
Eth1/6        RouterPort6   connected   routed    full 100G
QSFP-100G-CR4
Eth1/7        RouterPort7   connected   routed    full 100G
QSFP-100G-CR4
Eth1/8        celadm14      connected   3888      full 100G
QSFP-100G-CR4
Eth1/9        celadm13      connected   3888      full 100G
QSFP-100G-CR4
Eth1/10       celadm12      connected   3888      full 100G
QSFP-100G-CR4
Eth1/11       celadm11      connected   3888      full 100G
QSFP-100G-CR4
Eth1/12       celadm10      connected   3888      full 100G
QSFP-100G-CR4
  
```

```

Eth1/13      celadm09      connected 3888      full
100G      QSFP-100G-CR4
Eth1/14      celadm08      connected 3888      full
100G      QSFP-100G-CR4
Eth1/15      adm08         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/16      adm07         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/17      adm06         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/18      adm05         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/19      adm04         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/20      adm03         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/21      adm02         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/22      adm01         disabled  3888      full
100G      QSFP-100G-CR4
Eth1/23      celadm07      connected 3888      full
100G      QSFP-100G-CR4
Eth1/24      celadm06      connected 3888      full
100G      QSFP-100G-CR4
Eth1/25      celadm05      connected 3888      full
100G      QSFP-100G-CR4
Eth1/26      celadm04      connected 3888      full
100G      QSFP-100G-CR4
Eth1/27      celadm03      connected 3888      full
100G      QSFP-100G-CR4
Eth1/28      celadm02      connected 3888      full
100G      QSFP-100G-CR4
Eth1/29      celadm01      connected 3888      full
100G      QSFP-100G-CR4
Eth1/30      RouterPort8   connected routed      full
100G      QSFP-100G-CR4
Eth1/31      RouterPort9   connected routed      full
100G      QSFP-100G-CR4
Eth1/32      RouterPort10  connected routed      full
100G      QSFP-100G-CR4
Eth1/33      RouterPort11  connected routed      full
100G      QSFP-100G-CR4
Eth1/34      RouterPort12  connected routed      full
100G      QSFP-100G-CR4
Eth1/35      RouterPort13  connected routed      full
100G      QSFP-100G-CR4
Eth1/36      RouterPort14  connected routed      full
100G      QSFP-100G-CR4
Lo0          Routing loopback i connected routed      auto
auto        --
Lo1          VTEP loopback inte connected routed      auto
auto        --
Vlan1       --            down      routed      auto
auto        --
  
```

```
nve1          --          connected --          auto auto    --
```

- g. Start the database server ports on the upper leaf switches (R1UL and R2UL).

- i. On R1UL:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/15-22
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. Repeat the command sequence on R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/15-22
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

- h. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- i. Verify the status of the database server ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

- i. R1UL# **show interface status**

- ii. R2UL# **show interface status**

Examine the output to ensure that the database server ports are connected.

For example:

```
R1UL# show interface status

-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
-----
```

```

mgmt0      --          connected routed    full
1000      --
  
```

```

-----
Port          Name          Status  Vlan    Duplex
Speed   Type
-----
Eth1/1      RouterPort1   connected routed    full
100G      QSFP-100G-CR4
Eth1/2      RouterPort2   connected routed    full
100G      QSFP-100G-CR4
Eth1/3      RouterPort3   connected routed    full
100G      QSFP-100G-CR4
Eth1/4      RouterPort4   connected routed    full
100G      QSFP-100G-CR4
Eth1/5      RouterPort5   connected routed    full
100G      QSFP-100G-CR4
Eth1/6      RouterPort6   connected routed    full
100G      QSFP-100G-CR4
Eth1/7      RouterPort7   connected routed    full
100G      QSFP-100G-CR4
Eth1/8      celadm14      connected 3888    full
100G      QSFP-100G-CR4
Eth1/9      celadm13      connected 3888    full
100G      QSFP-100G-CR4
Eth1/10     celadm12      connected 3888    full
100G      QSFP-100G-CR4
Eth1/11     celadm11      connected 3888    full
100G      QSFP-100G-CR4
Eth1/12     celadm10      connected 3888    full
100G      QSFP-100G-CR4
Eth1/13     celadm09      connected 3888    full
100G      QSFP-100G-CR4
Eth1/14     celadm08      connected 3888    full
100G      QSFP-100G-CR4
Eth1/15     adm08         connected 3888    full
100G      QSFP-100G-CR4
Eth1/16     adm07         connected 3888    full
100G      QSFP-100G-CR4
Eth1/17     adm06         connected 3888    full
100G      QSFP-100G-CR4
Eth1/18     adm05         connected 3888    full
100G      QSFP-100G-CR4
Eth1/19     adm04         connected 3888    full
100G      QSFP-100G-CR4
Eth1/20     adm03         connected 3888    full
100G      QSFP-100G-CR4
Eth1/21     adm02         connected 3888    full
100G      QSFP-100G-CR4
Eth1/22     adm01         connected 3888    full
100G      QSFP-100G-CR4
Eth1/23     celadm07      connected 3888    full
100G      QSFP-100G-CR4
  
```

```

Eth1/24      celadm06      connected 3888      full 100G
QSFP-100G-CR4
Eth1/25      celadm05      connected 3888      full 100G
QSFP-100G-CR4
Eth1/26      celadm04      connected 3888      full 100G
QSFP-100G-CR4
Eth1/27      celadm03      connected 3888      full 100G
QSFP-100G-CR4
Eth1/28      celadm02      connected 3888      full 100G
QSFP-100G-CR4
Eth1/29      celadm01      connected 3888      full 100G
QSFP-100G-CR4
Eth1/30      RouterPort8   connected routed      full 100G
QSFP-100G-CR4
Eth1/31      RouterPort9   connected routed      full 100G
QSFP-100G-CR4
Eth1/32      RouterPort10  connected routed      full 100G
QSFP-100G-CR4
Eth1/33      RouterPort11  connected routed      full 100G
QSFP-100G-CR4
Eth1/34      RouterPort12  connected routed      full 100G
QSFP-100G-CR4
Eth1/35      RouterPort13  connected routed      full 100G
QSFP-100G-CR4
Eth1/36      RouterPort14  connected routed      full 100G
QSFP-100G-CR4
Lo0          Routing loopback i connected routed      auto auto      --
Lo1          VTEP loopback inte connected routed      auto auto      --
Vlan1       --            down      routed      auto auto      --
nve1        --            connected --          auto auto      --
  
```

j. Verify that nve is up on the leaf switches (R1LL, R1UL, R2LL, and R2UL).

Run the following command on each leaf switch and examine the output:

i. R1LL# **show nve peers**

ii. R1UL# **show nve peers**

iii. R2LL# **show nve peers**

iv. R2UL# **show nve peers**

In the output, you should see three nve peers with State=Up.

For example:

```

R1UL# show nve peers
Interface      Peer-IP State LearnType  Uptime      Router-Mac
-----
nve1          100.64.1.101 Up        CP 00:04:29 n/a
nve1          100.64.1.103 Up        CP 00:07:48 n/a
nve1          100.64.1.104 Up        CP 00:04:10 n/a
  
```

- k. Verify that BGP is up on the upper leaf switches (R1UL and R2UL).

Run the following command on each upper leaf switch and examine the output:

i. R1UL# **show logging log | grep BGP**

ii. R2UL# **show logging log | grep BGP**

In the output, look for two entries with `Up` in the rightmost column that are associated with different IP addresses.

For example:

```
R1UL# show logging log | grep BGP
2020 Aug 10 13:57:13 R1UL %BGP-5-ADJCHANGE: bgp- [32782]
(default) neighbor 100.64.0.201 Up
2020 Aug 10 13:57:24 R1UL %BGP-5-ADJCHANGE: bgp- [32782]
(default) neighbor 100.64.0.202 Up
```

7. For each rack (R1 and R2), confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

The `verify_roce_cables.py` script uses two input files; one for database servers and storage servers (`nodes.rackN`), and another for switches (`switches.rackN`). In each file, every server or switch must be listed on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

See My Oracle Support document 2587717.1 for download and detailed usage instructions.

Run the `verify_roce_cables.py` script against both of the racks:

a. # **cd /opt/oracle.SupportTools/ibdiagtools**
./verify_roce_cables.py -n nodes.rack1 -s switches.rack1

b. # **cd /opt/oracle.SupportTools/ibdiagtools**
./verify_roce_cables.py -n nodes.rack2 -s switches.rack2

Check the output of the `verify_roce_cables.py` script against the tables in [Two-Rack Cabling for X9M Racks](#). Also, check that output in the `CABLE OK?` columns contains the `OK` status.

The following examples show extracts of the expected command results:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER) LOWER LEAF (rack1sw-rocea0) : CABLE OK? UPPER LEAF
(rack1sw-roceb0) : CABLE OK?
----- : -----
----- : -----
...

# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
```

```
SWITCH PORT (EXPECTED PEER) LOWER LEAF (rack2sw-rocea0) : CABLE OK? UPPER LEAF
(rack2sw-roceb0) : CABLE OK?
----- : -----
----- : -----
...
```

8. Verify the RoCE Network Fabric operation across both interconnected racks by using the `infinicheck` command.

Use the following recommended command sequence to verify the RoCE Network Fabric operation across both racks.

In each command, `hosts.all` contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses from both racks (2 RoCE Network Fabric IP addresses for each database server), and `cells.all` contains a list of RoCE Network Fabric IP addresses for the storage servers from both racks (2 RoCE Network Fabric IP addresses for each storage server).

- a. `# cd /opt/oracle.SupportTools/ibdiagtools`
`# ./infinicheck -g hosts.all -c cells.all -z`
- b. `# cd /opt/oracle.SupportTools/ibdiagtools`
`# ./infinicheck -g hosts.all -c cells.all -s`
- c. `# cd /opt/oracle.SupportTools/ibdiagtools`
`# ./infinicheck -g hosts.all -c cells.all -b`

See step 1.k for most information about each `infinicheck` command.

The following example shows the expected command results for the final command in the sequence:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.all -c cells.all -b

INFINICHECK
      [Network Connectivity, Configuration and Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
      ##### ROCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS]... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS]... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS]... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs and CELLS
[SUCCESS]... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value : 2300
[SUCCESS]... RoCE interface MTU settings look good
      Verifying switch advertised DSCP on all DBs and CELLS ports ( )
[SUCCESS]... Advertised DSCP settings from RoCE switch looks good
      ##### CONNECTIVITY TESTS #####
```



```
[COMPUTE NODES -> STORAGE CELLS]
(60 seconds approx.)
(Will walk through QoS values: 0-6) [SUCCESS].....Results
OK
[SUCCESS]..... All can talk to all storage cells
[COMPUTE NODES -> COMPUTE NODES]
...
```

At this point, both racks share the RoCE Network Fabric, and the combined system is ready for further configuration. See [Configuring the New Hardware](#).

Related Topics

- [Cabling Two Oracle Exadata Database Machine RoCE Network Fabric Racks With No Downtime \(My Oracle Support Doc ID 2704997.1\)](#)
- [Verify RoCE Cabling on Oracle Exadata Database Machine X8M-2 and X8M-8 Servers \(My Oracle Support Doc ID 2587717.1\)](#)

2.3.2.1.2 Extending an X8M Rack with No Down Time by Adding an X9M Rack

WARNING:

Take time to read and understand this procedure before implementation. Pay careful attention to the instructions that surround the command examples. A system outage may occur if the procedure is not applied correctly.

Note:

This procedure assumes that the RoCE Network Fabric switches on the existing X8M rack contain the golden configuration settings from Oracle Exadata System Software 20.1.0 or later. Otherwise, before using this procedure, you must update the Oracle Exadata System Software and update the golden configuration settings on the RoCE Network Fabric switches. Downtime is required to update the golden configuration settings on the RoCE Network Fabric switches.

Note:

For additional background information, see [Understanding Multi-Rack Cabling for X8M Racks](#) and [Understanding Multi-Rack Cabling for X9M Racks](#).

Use this procedure to extend a typical X8M rack without down-time by cabling it together with an X9M rack. The primary rack (designated R1) and all of the systems it supports remain online throughout the procedure. At the beginning of the procedure, the additional rack (designated R2) is shut down.

The following is an outline of the procedure:

- Preparation (steps 1 and 2)

In this phase, you prepare the racks, switches, and cables. Also, you install and cable the spine switches in both racks.

- Configuration and Physical Cabling

In this phase, you reconfigure the leaf switches and finalize the cabling to the spine switches. These tasks are carefully orchestrated to avoid downtime on the primary system, as follows:

- Partially configure the lower leaf switches (step 3)

In this step, you reconfigure the switch ports on the lower leaf switches. There is no physical cabling performed in this step.

- Partially configure the upper leaf switches (step 4)

In this step, you reconfigure the switch ports on the upper leaf switches, remove the inter-switch cables that connect the leaf switches in both racks and connect the cables between the upper leaf switches and the spine switches.

- Finalize the lower leaf switches (step 5)

In this step, you finalize the switch port configuration on the lower leaf switches. You also complete the physical cabling by connecting the cables between the lower leaf switches and the spine switches.

- Finalize the upper leaf switches (step 6)

In this step, you finalize the switch port configuration on the upper leaf switches.

- Validation and Testing (steps 7 and 8)

In this phase, you validate and test the RoCE Network Fabric across both of the interconnect racks.

After completing the procedure, both racks share the RoCE Network Fabric, and the combined system is ready for further configuration. For example, you can extend existing disk groups and Oracle RAC databases to consume resources across both racks.

 **Note:**

- This procedure applies only to typical rack configurations that initially have leaf switches with the following specifications:
 - The inter-switch ports are ports 4 to 7, and ports 30 to 33.
 - The storage server ports are ports 8 to 14, and ports 23 to 29.
 - The database server ports are ports 15 to 22.
- For other rack configurations (for example, 8-socket systems with three database servers and 11 storage servers) a different procedure and different RoCE Network Fabric switch configuration files are required. Contact Oracle for further guidance.
- The procedure uses the following naming abbreviations and conventions:
 - The abbreviation for the existing X8M rack is R1, and the new X9M rack is R2.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies a spine switch.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the existing rack (R1).
 - Most operations must be executed in multiple locations. For example, step 1.h instructs you to update the firmware on all the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL). Pay attention to the instructions and keep track of your actions.

 **Tip:**

When a step must be performed on multiple switches, the instruction contains a list of the applicable switches. For example, (R1LL, R1UL, R2LL, and R2UL). You can use this list as a checklist to keep track of your actions.

- Perform operations sequentially, and complete every operation before proceeding. For example, run the entire command sequence at 3.a.i as one operation and complete it before proceeding.
 - All of commands that are run on a RoCE Network Fabric switch must be run while connected to the switch management interface as the switch administrator.
1. Prepare the systems.
 - a. Position the new X9M rack (R2) so that it is physically near the existing X8M rack (R1).

The RDMA Network Fabric cables must be able to reach the switches in each rack.

For the required cross-rack cabling information, see [Two-Rack Cabling for a System Combining X8M and X9M Racks](#).

- b. Power on all of the servers and network switches in the new rack (R2).
This includes the database servers, storage servers, RoCE Network Fabric leaf switches, and the Management Network Switch.
- c. Prepare the RoCE Network Fabric cables that you will use to interconnect the racks.
Label both ends of every cable.
For the required cross-rack cabling information, see [Two-Rack Cabling for a System Combining X8M and X9M Racks](#).
- d. Connect the new rack (R2) to your existing management network.
Ensure that there are no IP address conflicts across the racks and that you can access the management interfaces on the RoCE Network Fabric switches.
- e. Ensure that you have a backup of the current switch configuration for each RoCE Network Fabric switch (R1LL, R1UL, R2LL, and R2UL).
See [Backing Up Settings on the RoCE Network Fabric Switch in Oracle Exadata Database Machine Maintenance Guide](#).
- f. Download the required RoCE Network Fabric switch configuration files.
This procedure requires specific RoCE Network Fabric switch configuration files, which you must download from [My Oracle Support document 2704997.1](#).

 **WARNING:**

You must use different switch configuration files depending on whether your system uses Exadata Secure RDMA Fabric Isolation. Ensure that you download the correct archive that matches your system configuration.

For system configurations without Secure Fabric, download `online_multi-rack_8and14uplinks.zip`. For system configurations with Secure Fabric, download `online_SF_enabled_multi-rack_8and14uplinks.zip`.

Download and extract the archive containing the required RoCE Network Fabric switch configuration files. Place the files on a server with access to the management interfaces on the RoCE Network Fabric switches.

- g. Copy the required RoCE Network Fabric switch configuration files to the leaf switches on both racks.

You can use the following commands to copy the required configuration files to all of the RoCE Network Fabric switches on a system without Secure Fabric enabled:

- i.

```
# scp roce_multi_online_step3_R1_LL.cfg admin@R1LL_IP:/
```
- ii.

```
# scp roce_multi_14uplinks_online_step3_R2_LL.cfg admin@R2LL_IP:/
```
- iii.

```
# scp roce_multi_online_step4_R1_UL.cfg admin@R1UL_IP:/
```

- iv. # scp roce_multi_14uplinks_online_step4_R2_UL.cfg
admin@R2UL_IP:/
- v. # scp roce_multi_online_step5.cfg admin@R1LL_IP:/
- vi. # scp roce_multi_14uplinks_online_step5.cfg admin@R2LL_IP:/

On a system with Secure Fabric enabled, you can use the following commands:

- i. # scp roce_SF_multi_online_step3_R1_LL.cfg admin@R1LL_IP:/
- ii. # scp roce_SF_multi_14uplinks_online_step3_R2_LL.cfg
admin@R2LL_IP:/
- iii. # scp roce_SF_multi_online_step4_R1_UL.cfg admin@R1UL_IP:/
- iv. # scp roce_SF_multi_14uplinks_online_step4_R2_UL.cfg
admin@R2UL_IP:/
- v. # scp roce_SF_multi_online_step5.cfg admin@R1LL_IP:/
- vi. # scp roce_SF_multi_14uplinks_online_step5.cfg
admin@R2LL_IP:/

In the above commands, substitute the appropriate IP address or host name where applicable. For example, in place of *R1LL_IP*, substitute the management IP address or host name for the lower leaf switch (LL) on the existing rack (R1).

 **Note:**

The command examples in the rest of this procedure use the configuration files for a system configuration without Secure Fabric enabled. If required, adjust the commands to use the Secure Fabric-enabled switch configuration files.

- h. Update the firmware to the latest available release on all of the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL).

See *Updating RoCE Network Fabric Switch Firmware in Oracle Exadata Database Machine Maintenance Guide*.

- i. Examine the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL) and confirm the port categories for the cabled ports.

Run the `show interface status` command on every RoCE Network Fabric leaf switch:

- i. R1LL# `show interface status`

ii. R1U1# **show interface status**

iii. R2L1# **show interface status**

iv. R2U1# **show interface status**

Examine the output and confirm the port categories as follows:

- Confirm that the inter-switch ports are ports 4 to 7, and ports 30 to 33.
- Confirm that the storage server ports are ports 8 to 14, and ports 23 to 29.
- Confirm that the database server ports are ports 15 to 22.

For example:

R1L1# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
Eth1/1        --            xcvrAbsen  1          auto    auto    --
Eth1/2        --            xcvrAbsen  1          auto    auto    --
Eth1/3        --            xcvrAbsen  1          auto    auto    --
Eth1/4        ISL1          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/5        ISL2          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/6        ISL3          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/7        ISL4          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/8        celadm14     connected   3888     full    100G
QSFP-100G-CR4
Eth1/9        celadm13     connected   3888     full    100G
QSFP-100G-CR4
Eth1/10       celadm12     connected   3888     full    100G
QSFP-100G-CR4
Eth1/11       celadm11     connected   3888     full    100G
QSFP-100G-CR4
Eth1/12       celadm10     connected   3888     full    100G
QSFP-100G-CR4
Eth1/13       celadm09     connected   3888     full    100G
QSFP-100G-CR4
  
```

```

Eth1/14      celadm08      connected 3888      full
100G      QSFP-100G-CR4
Eth1/15      adm08        connected 3888      full
100G      QSFP-100G-CR4
Eth1/16      adm07        connected 3888      full
100G      QSFP-100G-CR4
Eth1/17      adm06        connected 3888      full
100G      QSFP-100G-CR4
Eth1/18      adm05        connected 3888      full
100G      QSFP-100G-CR4
Eth1/19      adm04        connected 3888      full
100G      QSFP-100G-CR4
Eth1/20      adm03        connected 3888      full
100G      QSFP-100G-CR4
Eth1/21      adm02        connected 3888      full
100G      QSFP-100G-CR4
Eth1/22      adm01        connected 3888      full
100G      QSFP-100G-CR4
Eth1/23      celadm07     connected 3888      full
100G      QSFP-100G-CR4
Eth1/24      celadm06     connected 3888      full
100G      QSFP-100G-CR4
Eth1/25      celadm05     connected 3888      full
100G      QSFP-100G-CR4
Eth1/26      celadm04     connected 3888      full
100G      QSFP-100G-CR4
Eth1/27      celadm03     connected 3888      full
100G      QSFP-100G-CR4
Eth1/28      celadm02     connected 3888      full
100G      QSFP-100G-CR4
Eth1/29      celadm01     connected 3888      full
100G      QSFP-100G-CR4
Eth1/30      ISL5         connected trunk    full
100G      QSFP-100G-CR4
Eth1/31      ISL6         connected trunk    full
100G      QSFP-100G-CR4
Eth1/32      ISL7         connected trunk    full
100G      QSFP-100G-CR4
Eth1/33      ISL8         connected trunk    full
100G      QSFP-100G-CR4
Eth1/34      --           xcvrAbsen 1       auto
auto      --
Eth1/35      --           xcvrAbsen 1       auto
auto      --
Eth1/36      --           xcvrAbsen 1       auto
auto      --
Po100      --           connected trunk    full
100G      --
Lo0        Routing loopback i connected routed    auto
auto      --
Lo1        VTEP loopback inte connected routed    auto
auto      --
Vlan1      --           down      routed    auto
auto      --

```

```
nvel          --          connected --          auto    auto    --
```

- j. For each rack (R1 and R2), confirm the RoCE Network Fabric cabling by running the `verify_roce_cables.py` script.

The `verify_roce_cables.py` script uses two input files; one for database servers and storage servers (`nodes.rackN`), and another for switches (`switches.rackN`). In each file, every server or switch must be listed on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

See My Oracle Support document 2587717.1 for download and detailed usage instructions.

Run the `verify_roce_cables.py` script against both of the racks:

- i.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
```
- ii.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
```

Check that output in the `CABLE OK?` columns contains the `OK` status.

The following example shows the expected command results:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER)          LOWER LEAF (rack1sw-
rocea0) : CABLE OK?                   UPPER LEAF (rack1sw-roceb0) : CABLE
OK?
-----
----- : -----
----- : -----
Eth1/4 (ISL peer switch) : rack1sw-rocea0
Ethernet1/4 : OK                rack1sw-roceb0 Ethernet1/4 : OK
Eth1/5 (ISL peer switch) : rack1sw-rocea0
Ethernet1/5 : OK                rack1sw-roceb0 Ethernet1/5 : OK
Eth1/6 (ISL peer switch) : rack1sw-rocea0
Ethernet1/6 : OK                rack1sw-roceb0 Ethernet1/6 : OK
Eth1/7 (ISL peer switch) : rack1sw-rocea0
Ethernet1/7 : OK                rack1sw-roceb0 Ethernet1/7 : OK
Eth1/8 (RU39)             : rack1celadm14
port-1 : OK                    rack1celadm14 port-2 : OK
Eth1/9 (RU37)             : rack1celadm13
port-1 : OK                    rack1celadm13 port-2 : OK
Eth1/10 (RU35)            : rack1celadm12
port-1 : OK                    rack1celadm12 port-2 : OK
Eth1/11 (RU33)            : rack1celadm11
port-1 : OK                    rack1celadm11 port-2 : OK
Eth1/12 (RU31)            : rack1celadm10
port-1 : OK                    rack1celadm10 port-2 : OK
Eth1/13 (RU29)            : rack1celadm09
port-1 : OK                    rack1celadm09 port-2 : OK
Eth1/14 (RU27)            : rack1celadm08
port-1 : OK                    rack1celadm08 port-2 : OK
Eth1/15 (RU26)            : rack1ladm08
```



```

port-1 : OK
  Eth1/16 (RU25) : rack1adm8 port-2 : OK
port-1 : OK
  Eth1/17 (RU24) : rack1adm7 port-2 : OK
port-1 : OK
  Eth1/18 (RU23) : rack1adm6 port-2 : OK
port-1 : OK
  Eth1/19 (RU19) : rack1adm5 port-2 : OK
port-1 : OK
  Eth1/20 (RU18) : rack1adm4 port-2 : OK
port-1 : OK
  Eth1/21 (RU17) : rack1adm3 port-2 : OK
port-1 : OK
  Eth1/22 (RU16) : rack1adm2 port-2 : OK
port-1 : OK
  Eth1/23 (RU14) : rack1adm1 port-2 : OK
port-1 : OK
  Eth1/24 (RU12) : rack1celadm7 port-2 : OK
port-1 : OK
  Eth1/25 (RU10) : rack1celadm6 port-2 : OK
port-1 : OK
  Eth1/26 (RU08) : rack1celadm5 port-2 : OK
port-1 : OK
  Eth1/27 (RU06) : rack1celadm4 port-2 : OK
port-1 : OK
  Eth1/28 (RU04) : rack1celadm3 port-2 : OK
port-1 : OK
  Eth1/29 (RU02) : rack1celadm2 port-2 : OK
port-1 : OK
  Eth1/30 (ISL peer switch) : rack1celadm1 port-2 : OK
Ethernet1/30 : OK rack1sw-rocea0
Ethernet1/30 : OK rack1sw-roceb0
  Eth1/31 (ISL peer switch) : rack1sw-rocea0
Ethernet1/31 : OK rack1sw-roceb0
Ethernet1/31 : OK
  Eth1/32 (ISL peer switch) : rack1sw-rocea0
Ethernet1/32 : OK rack1sw-roceb0
Ethernet1/32 : OK
  Eth1/33 (ISL peer switch) : rack1sw-rocea0
Ethernet1/33 : OK rack1sw-roceb0
Ethernet1/33 : OK
  
```

- k. For each rack (R1 and R2), verify the RoCE Network Fabric operation by using the `infinicheck` command.
- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command.
 - Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric.
 - Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks.

In each command, the hosts input file (`hosts.rack1` and `hosts.rack2`) contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses (2 RoCE Network Fabric IP addresses for each database server), and the cells input file (`cells.rack1` and `cells.rack2`) contains a list of RoCE Network Fabric IP addresses for the storage servers (2 RoCE Network Fabric IP addresses for each storage server).

i. Use the following recommended command sequence on the existing rack (R1):

```
i. # cd /opt/oracle.SupportTools/ibdiagtools
   # ./infinicheck -g hosts.rack1 -c cells.rack1 -z

ii. # cd /opt/oracle.SupportTools/ibdiagtools
    # ./infinicheck -g hosts.rack1 -c cells.rack1 -s

iii. # cd /opt/oracle.SupportTools/ibdiagtools
     # ./infinicheck -g hosts.rack1 -c cells.rack1 -b
```

ii. Use the following recommended command sequence on the new rack (R2):

```
i. # cd /opt/oracle.SupportTools/ibdiagtools
   # ./infinicheck -g hosts.rack2 -c cells.rack2 -z

ii. # cd /opt/oracle.SupportTools/ibdiagtools
    # ./infinicheck -g hosts.rack2 -c cells.rack2 -s

iii. # cd /opt/oracle.SupportTools/ibdiagtools
     # ./infinicheck -g hosts.rack2 -c cells.rack2 -b
```

The following example shows the expected command results for the final command in the sequence:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rackN -c cells.rackN -b

INFINICHECK
      [Network Connectivity, Configuration and Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS].... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS].... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS].... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs and
CELLS
[SUCCESS].... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value : 2300
[SUCCESS].... RoCE interface MTU settings look good
```

```

    Verifying switch advertised DSCP on all DBs and CELLS ports
  ( )
[SUCCESS].... Advertised DSCP settings from RoCE switch looks
good
    ##### CONNECTIVITY TESTS #####
    [COMPUTE NODES -> STORAGE CELLS]
    (60 seconds approx.)
    (Will walk through QoS values: 0-6)
[SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
    [COMPUTE NODES -> COMPUTE NODES]
  ...
  
```

2. Install the spine switches (R1SS and R2SS).

- a. Physically install and power up the spine switches in the existing rack (R1SS) and the new rack (R2SS).
 - i. Physically install each spine switch in RU1.
 - ii. For each spine switch, ensure that the management Ethernet interface is connected to the management network and then supply power.
 - iii. On each spine switch, perform the initial configuration steps outlined in Configuring the Cisco Nexus C9336C-FX2 Switch. Skip the step for applying the golden configuration settings as you will do this later.
 - iv. For each spine switch, perform a ping test to the management Ethernet interface to ensure that the switch is online and accessible.

b. Apply the golden configuration settings to the new spine switches.

See Applying Golden Configuration Settings on RoCE Network Fabric Switches in *Oracle Exadata Database Machine Maintenance Guide*.

You can use the instance of `patchmgr` that you previously used to update the firmware on the leaf switches (in step 1.h).

Use a switch list file (`spines.lst`) to apply the golden configuration settings to both spine switches using one `patchmgr` command:

```

# cat spines.lst
R1SS_IP:m spine.201
R2SS_IP:m spine.202

# ./patchmgr --roceswitches spines.lst --apply-config -
log_dir /tmp/spinelogs
  
```

 **Note:**

In the switch list file, `R1SS_IP` is the management IP address or host name for the spine switch on the existing rack (R1SS) and `R2SS_IP` is the management IP address or host name for the spine switch on the new rack (R2SS).

c. Upgrade the firmware on the spine switches.

See Updating RoCE Network Fabric Switch Firmware in *Oracle Exadata Database Machine Maintenance Guide*.

You can use the instance of `patchmgr` that you used in the previous step.

Use a switch list file (`spines.lst`) to perform the firmware upgrade on both spine switches using one `patchmgr` command:

```
# cat spines.lst
R1SS_IP:mspine.201
R2SS_IP:mspine.202

# ./patchmgr --roceswitches spines.lst --upgrade -log_dir /tmp/
spinelogs
```

 **Note:**

In the switch list file, `R1SS_IP` is the management IP address or host name for the spine switch on the existing rack (R1SS) and `R2SS_IP` is the management IP address or host name for the spine switch on the new rack (R2SS).

- d. Connect the RoCE Network Fabric cables to the spine switches (R1SS and R2SS).

 **WARNING:**

**At this stage, only connect the cables to the spine switches.
DO NOT CONNECT ANY OF THE CABLES TO THE LEAF SWITCHES.**

Use the cabled that you prepared earlier (in step 1.c).

For the required cross-rack cabling information, see [Two-Rack Cabling for a System Combining X8M and X9M Racks](#).

3. Perform the first round of configuration on the lower leaf switches (R1LL and R2LL).

Perform this step on the lower leaf switches (R1LL and R2LL) only.

 **Note:**

During this step, the lower leaf switch ports are shut down. While the R1LL ports are down, R1UL exclusively supports the RoCE Network Fabric. During this time, there is no redundancy in the RoCE Network Fabric, and availability cannot be maintained if R1UL goes down.

- a. Shut down the switch ports on the lower leaf switches (R1LL and R2LL).

i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/4-33
R1LL(config-if-range)# shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. On R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/1-36
R2LL(config-if-range)# shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

b. Reconfigure the lower leaf switch ports (R1LL and R2LL) .

For each switch, you must use the correct corresponding switch configuration file, which you earlier copied to the switch (in step 1.g).

i. On R1LL, the switch configuration file name must end with

step3_R1_LL.cfg:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# run-script bootflash:///
roce_multi_online_step3_R1_LL.cfg | grep 'none'
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. On R2LL, the switch configuration file name must end with

step3_R2_LL.cfg:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step3_R2_LL.cfg | grep 'none'
R2LL(config)# copy running-config startup-config
```

```
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- c. Start the inter-switch ports on the lower leaf switches (R1LL and R2LL) .

- i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/4-7, ethernet 1/30-33
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. On R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/1-7, ethernet 1/30-36
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- d. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- e. Verify the status of the inter-switch ports on the lower leaf switches (R1LL and R2LL) .

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**
- ii. R2LL# **show interface status**

Examine the output to ensure that the inter-switch ports are connected.

For example:

R1LL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex
Speed   Type
-----
mgmt0         --           connected  routed    full
1000         --
  
```

```

-----
-----
Port          Name          Status      Vlan      Duplex
Speed   Type
-----
Eth1/1         --           xcvrAbsen  1         auto
auto          --
Eth1/2         --           xcvrAbsen  1         auto
auto          --
Eth1/3         --           xcvrAbsen  1         auto
auto          --
Eth1/4         ISL1         connected  trunk     full
100G          QSF-100G-CR4
Eth1/5         ISL2         connected  trunk     full
100G          QSF-100G-CR4
Eth1/6         ISL3         connected  trunk     full
100G          QSF-100G-CR4
Eth1/7         ISL4         connected  trunk     full
100G          QSF-100G-CR4
Eth1/8         celadm14    disabled   3888     full
100G          QSF-100G-CR4
Eth1/9         celadm13    disabled   3888     full
100G          QSF-100G-CR4
Eth1/10        celadm12    disabled   3888     full
100G          QSF-100G-CR4
Eth1/11        celadm11    disabled   3888     full
100G          QSF-100G-CR4
Eth1/12        celadm10    disabled   3888     full
100G          QSF-100G-CR4
Eth1/13        celadm09    disabled   3888     full
100G          QSF-100G-CR4
Eth1/14        celadm08    disabled   3888     full
100G          QSF-100G-CR4
Eth1/15        adm08       disabled   3888     full
100G          QSF-100G-CR4
Eth1/16        adm07       disabled   3888     full
100G          QSF-100G-CR4
Eth1/17        adm06       disabled   3888     full
100G          QSF-100G-CR4
Eth1/18        adm05       disabled   3888     full
100G          QSF-100G-CR4
  
```

Eth1/19	adm04	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/20	adm03	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/21	adm02	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/22	adm01	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/23	celadm07	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/24	celadm06	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/25	celadm05	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/26	celadm04	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/27	celadm03	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/28	celadm02	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/29	celadm01	disabled	3888	full	100G		
QSFP-100G-CR4							
Eth1/30	ISL5	connected	trunk	full	100G		
QSFP-100G-CR4							
Eth1/31	ISL6	connected	trunk	full	100G		
QSFP-100G-CR4							
Eth1/32	ISL7	connected	trunk	full	100G		
QSFP-100G-CR4							
Eth1/33	ISL8	connected	trunk	full	100G		
QSFP-100G-CR4							
Eth1/34	--	xcvrAbsen	1	auto	auto	--	
Eth1/35	--	xcvrAbsen	1	auto	auto	--	
Eth1/36	--	xcvrAbsen	1	auto	auto	--	
Po100	--	connected	trunk	full	100G	--	
Lo0	Routing loopback	i	connected	routed	auto	auto	--
Lo1	VTEP loopback	inte	connected	routed	auto	auto	--
Vlan1	--	down	routed	auto	auto	--	
nve1	--	connected	--	auto	auto	--	

f. Start the storage server ports on the lower leaf switches (R1LL and R2LL) .

i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```


- ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- g. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- h. Verify the status of the storage server ports on the lower leaf switches (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**
- ii. R2LL# **show interface status**

Examine the output to ensure that the storage server ports are connected.

For example:

```
R1LL# show interface status
```

Port	Name	Status	Vlan	Duplex
mgmt0	--	connected	routed	full
1000	--			

Port	Name	Status	Vlan	Duplex
Eth1/1	--	xcvrAbsen	1	auto
auto	--			
Eth1/2	--	xcvrAbsen	1	auto
auto	--			
Eth1/3	--	xcvrAbsen	1	auto
auto	--			
Eth1/4	ISL1	connected	trunk	full
100G	QSFP-100G-CR4			
Eth1/5	ISL2	connected	trunk	full

100G	QSFP-100G-CR4				
Eth1/6	ISL3	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/7	ISL4	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/8	celadm14	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/9	celadm13	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/10	celadm12	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/11	celadm11	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/12	celadm10	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/13	celadm09	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/14	celadm08	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/15	adm08	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/16	adm07	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/17	adm06	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/18	adm05	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/19	adm04	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/20	adm03	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/21	adm02	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/22	adm01	disabled 3888	full	100G	
QSFP-100G-CR4					
Eth1/23	celadm07	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/24	celadm06	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/25	celadm05	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/26	celadm04	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/27	celadm03	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/28	celadm02	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/29	celadm01	<i>connected</i> 3888	full	100G	
QSFP-100G-CR4					
Eth1/30	ISL5	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/31	ISL6	connected trunk	full	100G	
QSFP-100G-CR4					
Eth1/32	ISL7	connected trunk	full	100G	
QSFP-100G-CR4					

```

Eth1/33      ISL8          connected trunk    full
100G      QSFP-100G-CR4
Eth1/34      --                xcvrAbsen 1      auto
auto      --
Eth1/35      --                xcvrAbsen 1      auto
auto      --
Eth1/36      --                xcvrAbsen 1      auto
auto      --
Po100       --                connected trunk    full
100G       --
Lo0         Routing loopback i connected routed    auto
auto      --
Lo1         VTEP loopback inte connected routed    auto
auto      --
Vlan1       --                down      routed    auto
auto      --
nve1        --                connected --      auto
auto      --
  
```

- i. Start the database server ports on the lower leaf switches (R1LL and R2LL).

- i. On R1LL:

```

R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/15-22
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
  
```

- ii. Repeat the command sequence on R2LL:

```

R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/15-22
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
  
```

- j. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- k. Verify the status of the database server ports on the lower leaf switches (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

i. R1LL# **show interface status**

ii. R2LL# **show interface status**

Examine the output to ensure that the database server ports are connected.

For example:

R1LL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
Eth1/1        --            xcvrAbsen  1         auto    auto    --
Eth1/2        --            xcvrAbsen  1         auto    auto    --
Eth1/3        --            xcvrAbsen  1         auto    auto    --
Eth1/4        ISL1          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/5        ISL2          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/6        ISL3          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/7        ISL4          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/8        celadm14     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/9        celadm13     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/10       celadm12     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/11       celadm11     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/12       celadm10     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/13       celadm09     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/14       celadm08     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/15       adm08        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/16       adm07        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/17       adm06        connected   3888     full    100G

```

```

QSF1P-100G-CR4
Eth1/18      adm05      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/19      adm04      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/20      adm03      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/21      adm02      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/22      adm01      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/23      celadm07   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/24      celadm06   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/25      celadm05   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/26      celadm04   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/27      celadm03   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/28      celadm02   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/29      celadm01   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/30      ISL5       connected trunk    full
100G      QSF1P-100G-CR4
Eth1/31      ISL6       connected trunk    full
100G      QSF1P-100G-CR4
Eth1/32      ISL7       connected trunk    full
100G      QSF1P-100G-CR4
Eth1/33      ISL8       connected trunk    full
100G      QSF1P-100G-CR4
Eth1/34      --         xcvrAbsen 1       auto
auto      --
Eth1/35      --         xcvrAbsen 1       auto
auto      --
Eth1/36      --         xcvrAbsen 1       auto
auto      --
Po100      --         connected trunk    full
100G      --
Lo0        Routing loopback i connected routed    auto
auto      --
Lo1        VTEP loopback inte connected routed    auto
auto      --
Vlan1     --         down      routed    auto
auto      --
nve1      --         connected --      auto
auto      --
  
```

 **Note:**

Before proceeding, ensure that you have completed all of the actions in step 3 on both lower leaf switches (R1LL and R2LL). If not, then ensure that you go back and perform the missing actions.

4. Perform the first round of configuration on the upper leaf switches (R1UL and R2UL).
Perform this step on the upper leaf switches (R1UL and R2UL) only.

 **Note:**

At the start of this step, the upper leaf switch ports are shut down. While the R1UL ports are down, R1LL exclusively supports the RoCE Network Fabric on the existing rack. During this time, there is no redundancy in the RoCE Network Fabric, and availability cannot be maintained if R1LL goes down.

- a. Shut down the upper leaf switch ports (R1UL and R2UL).

- i. On R1UL:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/4-33
R1UL(config-if-range)# shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. On R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/1-36
R2UL(config-if-range)# shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

- b. On both racks, remove the inter-switch links between the leaf switches (R1LL to R1UL, and R2LL to R2UL).

On every leaf switch, remove the cables for the inter-switch links:

- i. On R1LL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
 - ii. On R1UL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
 - iii. On R2LL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
 - iv. On R2UL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
- c. On both racks, cable the upper leaf switch to both of the spine switches (R1UL and R2UL to R1SS and R2SS).

Connect the cables from the spine switches that you prepared earlier (in step 2.d).

Cable the switches as described in [Two-Rack Cabling for a System Combining X8M and X9M Racks](#):

- i. On R1UL, cable ports 04, 05, 06, 07, 30, 31, 32, and 33 to R1SS and R2SS.
 - ii. On R2UL, cable ports 01, 02, 03, 04, 05, 06, 07, 30, 31, 32, 33, 34, 35, and 36 to R1SS and R2SS.
- d. Reconfigure the upper leaf switch ports (R1UL and R2UL).

For each switch, you must use the correct corresponding switch configuration file, which you earlier copied to the switch (in step 1.g):

- i. On R1UL, the switch configuration file name must end with `step4_R1_UL.cfg`:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# run-script bootflash:///
roce_multi_online_step4_R1_UL.cfg | grep 'none'
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. On R2UL, the switch configuration file name must end with `step4_R2_UL.cfg`:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step4_R2_UL.cfg | grep 'none'
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- e. Check the status of the RoCE Network Fabric ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

i. R1UL# **show interface status**

ii. R2UL# **show interface status**

Examine the output to ensure that all of the cabled ports are `disabled`.

The following example shows the expected output on the X9M rack (R2UL). On the X8M rack (R1UL), ports 01, 02, 03, 34, 35, and 36 are not physically connected.

R2UL# **show interface status**

```

-----
-----
Port          Name                Status   Vlan    Duplex Speed
Type
-----
-----
mgmt0         --                  connected routed   full 1000  --
-----
-----
Port          Name                Status   Vlan    Duplex Speed
Type
-----
-----
Eth1/1       RouterPort1         disabled routed   full 100G
QSFP-100G-CR4
Eth1/2       RouterPort2         disabled routed   full 100G
QSFP-100G-CR4
Eth1/3       RouterPort3         disabled routed   full 100G
QSFP-100G-CR4
Eth1/4       RouterPort4         disabled routed   full 100G
QSFP-100G-CR4
Eth1/5       RouterPort5         disabled routed   full 100G
QSFP-100G-CR4
Eth1/6       RouterPort6         disabled routed   full 100G
QSFP-100G-CR4
Eth1/7       RouterPort7         disabled routed   full 100G
QSFP-100G-CR4
Eth1/8       celadm14            disabled 3888    full 100G
QSFP-100G-CR4
Eth1/9       celadm13            disabled 3888    full 100G
QSFP-100G-CR4
Eth1/10      celadm12            disabled 3888    full 100G
QSFP-100G-CR4
  
```



```

Eth1/11      celadm11      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/12      celadm10      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/13      celadm09      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/14      celadm08      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/15      adm08         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/16      adm07         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/17      adm06         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/18      adm05         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/19      adm04         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/20      adm03         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/21      adm02         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/22      adm01         disabled 3888      full
100G      QSFP-100G-CR4
Eth1/23      celadm07      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/24      celadm06      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/25      celadm05      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/26      celadm04      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/27      celadm03      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/28      celadm02      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/29      celadm01      disabled 3888      full
100G      QSFP-100G-CR4
Eth1/30      RouterPort8   disabled routed      full
100G      QSFP-100G-CR4
Eth1/31      RouterPort9   disabled routed      full
100G      QSFP-100G-CR4
Eth1/32      RouterPort10  disabled routed      full
100G      QSFP-100G-CR4
Eth1/33      RouterPort11  disabled routed      full
100G      QSFP-100G-CR4
Eth1/34      RouterPort12  disabled routed      full
100G      QSFP-100G-CR4
Eth1/35      RouterPort13  disabled routed      full
100G      QSFP-100G-CR4
Eth1/36      RouterPort14  disabled routed      full
100G      QSFP-100G-CR4
Lo0          Routing loopback i connected routed      auto
auto        --
Lo1          VTEP loopback inte connected routed      auto

```

```

auto      --
Vlan1     --                down      routed   auto auto  --
nve1      --                connected --        auto auto  --

```

 **Note:**

Before proceeding, ensure that you have completed all of the actions to this point in step 4 on both upper leaf switches (R1UL and R2UL). If not, then ensure that you go back and perform the missing actions.

f. Verify the configuration of the upper leaf switches.

You can use the instance of `patchmgr` that you previously used to update the switch firmware (in step 1.h).

Use a switch list file (`ul.lst`) to check both upper leaf switches using one `patchmgr` command:

```

# cat ul.lst
R1UL_IP:mleaf.102
R2UL_IP:mleaf_u14.104

```

On a system with Secure Fabric enabled, use the `msfleaf` and `msfleaf_u14` tags in the switch list file:

```

# cat ul.lst
R1UL_IP:msfleaf.102
R2UL_IP:msfleaf_u14.104

```

The following shows the recommended command and an example of the expected results:

```

# ./patchmgr --roceswitches ul.lst --verify-config -log_dir /tmp/log
2020-08-10 13:40:09 -0700           :Working: Initiating config
verification... Expect up to 6 minutes for each switch
Mon Aug 10 13:40:13 PDT 2020 1 of 4 :Verifying config on switch ...
...
Mon Aug 10 13:40:32 PDT 2020:      [INFO      ] Config matches
template: ...
Mon Aug 10 13:40:32 PDT 2020:      [SUCCESS   ] Config validation
successful!
2020-08-10 13:40:32 -0700
  Config check on RoCE switch(es)
2020-08-10 13:40:32 -0700
  Completed run of command: ./patchmgr --roceswitches ul.lst --verify-
config -log_dir /tmp/log
2020-08-10 13:40:32 -0700           :INFO      : config attempted on nodes
in file ul.lst: [R1UL_IP R2UL_IP]
2020-08-10 13:40:32 -0700           :INFO      : For details, check the
following files in /tmp/log:
2020-08-10 13:40:32 -0700           :INFO      : - updateRoceSwitch.log
2020-08-10 13:40:32 -0700           :INFO      : - updateRoceSwitch.trc

```

```
2020-08-10 13:40:32 -0700      :INFO   : - patchmgr.stdout
2020-08-10 13:40:32 -0700      :INFO   : - patchmgr.stderr
2020-08-10 13:40:32 -0700      :INFO   : - patchmgr.log
2020-08-10 13:40:32 -0700      :INFO   : - patchmgr.trc
2020-08-10 13:40:32 -0700      :INFO   : Exit status:0
2020-08-10 13:40:32 -0700      :INFO   : Exiting.
```

In the command output, verify that the switch configuration is good for both upper leaf switches. You can ignore messages about the ports that are down.

5. Finalize the configuration of the lower leaf switches (R1LL and R2LL).

Perform this step on the lower leaf switches (R1LL and R2LL) only.

a. Reconfigure the lower leaf switch ports (R1LL and R2LL).

Run the following command sequence on both of the lower leaf switches (R1LL and R2LL).

You must use the correct switch configuration file, which you earlier copied to the switch (in step 1.g). In this step, the configuration file name must end with step5.cfg.

i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# run-script bootflash:///
roce_multi_online_step5.cfg | grep 'none'
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. On R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# run-script bootflash:///
roce_multi_14uplinks_online_step5.cfg | grep 'none'
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

b. On both racks, cable the lower leaf switch to both of the spine switches (R1LL and R2LL to R1SS and R2SS).

Connect the cables from the spine switches that you prepared earlier (in step 2.d).

Cable the switches as described in [Two-Rack Cabling for a System Combining X8M and X9M Racks](#):

- i. On R1LL, cable ports 04, 05, 06, 07, 30, 31, 32, and 33 to R1SS and R2SS.
 - ii. On R2LL, cable ports 01, 02, 03, 04, 05, 06, 07, 30, 31, 32, 33, 34, 35, and 36 to R1SS and R2SS.
- c. On the lower leaf switches, verify that all of the cabled RoCE Network Fabric ports are connected (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

i. R1LL# `show interface status`

ii. R2LL# `show interface status`

Examine the output to ensure that all of the cabled ports are `connected`.

The following example shows the expected output on the X9M rack (R2LL). On the X8M rack (R1LL), ports 01, 02, 03, 34, 35, and 36 are not physically connected.

R2LL# `show interface status`

```

-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
mgmt0         --            connected   routed    full 1000  --
-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
Eth1/1       RouterPort1   connected   routed    full 100G
QSFP-100G-CR4
Eth1/2       RouterPort2   connected   routed    full 100G
QSFP-100G-CR4
Eth1/3       RouterPort3   connected   routed    full 100G
QSFP-100G-CR4
Eth1/4       RouterPort4   connected   routed    full 100G
QSFP-100G-CR4
Eth1/5       RouterPort5   connected   routed    full 100G
QSFP-100G-CR4
Eth1/6       RouterPort6   connected   routed    full 100G
QSFP-100G-CR4
Eth1/7       RouterPort7   connected   routed    full 100G
QSFP-100G-CR4
Eth1/8       celadm14      connected   3888      full 100G
QSFP-100G-CR4
Eth1/9       celadm13      connected   3888      full 100G

```

```

QSF1P-100G-CR4
Eth1/10      celadm12      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/11      celadm11      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/12      celadm10      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/13      celadm09      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/14      celadm08      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/15      adm08         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/16      adm07         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/17      adm06         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/18      adm05         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/19      adm04         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/20      adm03         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/21      adm02         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/22      adm01         connected 3888      full
100G        QSF1P-100G-CR4
Eth1/23      celadm07      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/24      celadm06      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/25      celadm05      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/26      celadm04      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/27      celadm03      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/28      celadm02      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/29      celadm01      connected 3888      full
100G        QSF1P-100G-CR4
Eth1/30      RouterPort8   connected routed    full
100G        QSF1P-100G-CR4
Eth1/31      RouterPort9   connected routed    full
100G        QSF1P-100G-CR4
Eth1/32      RouterPort10  connected routed    full
100G        QSF1P-100G-CR4
Eth1/33      RouterPort11  connected routed    full
100G        QSF1P-100G-CR4
Eth1/34      RouterPort12  connected routed    full
100G        QSF1P-100G-CR4
Eth1/35      RouterPort13  connected routed    full
100G        QSF1P-100G-CR4
Eth1/36      RouterPort14  connected routed    full
100G        QSF1P-100G-CR4
  
```

```

Lo0          Routing loopback i connected routed    auto auto    --
Lo1          VTEP loopback inte connected routed    auto auto    --
Vlan1       --                               down    routed    auto auto    --
nve1        --                               connected --      auto auto    --

```

 **Note:**

Before proceeding, ensure that you have completed all of the actions to this point in step 5 on both lower leaf switches (R1LL and R2LL). If not, then ensure that you go back and perform the missing actions.

d. Verify the configuration of the lower leaf switches.

You can use the instance of `patchmgr` that you previously used to update the switch firmware (in step 1.h).

Use a switch list file (`ll.lst`) to check both lower leaf switches using one `patchmgr` command:

```

# cat ll.lst
R1LL_IP:mleaf.101
R2LL_IP:mleaf_u14.103

```

On a system with Secure Fabric enabled, use the `msfleaf` and `msfleaf_u14` tags in the switch list file:

```

# cat ll.lst
R1LL_IP:msfleaf.101
R2LL_IP:msfleaf_u14.103

```

The following shows the recommended command and an example of the expected results:

```

# ./patchmgr --roceswitches ll.lst --verify-config -log_dir /tmp/log
2020-08-10 13:45:09 -0700          :Working: Initiating config
verification... Expect up to 6 minutes for each switch
Mon Aug 10 13:45:13 PDT 2020 1 of 4 :Verifying config on switch ...
...
Mon Aug 10 13:45:32 PDT 2020:      [INFO      ] Config matches
template: ...
Mon Aug 10 13:45:32 PDT 2020:      [SUCCESS   ] Config validation
successful!
2020-08-10 13:45:32 -0700
  Config check on RoCE switch(es)
2020-08-10 13:45:32 -0700
  Completed run of command: ./patchmgr --roceswitches ll.lst --verify-
config -log_dir /tmp/log
2020-08-10 13:45:32 -0700          :INFO      : config attempted on nodes
in file ll.lst: [R1LL_IP R2LL_IP]
2020-08-10 13:45:32 -0700          :INFO      : For details, check the
following files in /tmp/log:
2020-08-10 13:45:32 -0700          :INFO      : - updateRoceSwitch.log

```

```

2020-08-10 13:45:32 -0700      :INFO   :   -
updateRoceSwitch.trc
2020-08-10 13:45:32 -0700      :INFO   :   - patchmgr.stdout
2020-08-10 13:45:32 -0700      :INFO   :   - patchmgr.stderr
2020-08-10 13:45:32 -0700      :INFO   :   - patchmgr.log
2020-08-10 13:45:32 -0700      :INFO   :   - patchmgr.trc
2020-08-10 13:45:32 -0700      :INFO   : Exit status:0
2020-08-10 13:45:32 -0700      :INFO   : Exiting.
  
```

In the command output, verify that the switch configuration is good for both lower leaf switches.

- e. Verify that nve is up on the lower leaf switches (R1LL and R2LL).

Run the following command on each lower leaf switch and examine the output:

- i. R1LL# **show nve peers**

- ii. R2LL# **show nve peers**

At this point, you should see one nve peer with State=Up.

For example:

```

R1LL# show nve peers
Interface      Peer-IP State LearnType  Uptime      Router-
Mac
-----
-----
nve1           100.64.1.103  Up        CP 00:04:29
n/a
  
```

- f. Verify that BGP is up on the lower leaf switches (R1LL and R2LL).

Run the following command on each lower leaf switch and examine the output:

- i. R1LL# **show logging log | grep BGP**

- ii. R2LL# **show logging log | grep BGP**

Look for two entries with Up in the rightmost column that are associated with different IP addresses.

For example:

```

R1LL# show logging log | grep BGP
2020 Aug 10 13:47:13 R1LL %BGP-5-ADJCHANGE: bgp- [29342]
(default) neighbor 100.64.0.201 Up
2020 Aug 10 13:47:24 R1LL %BGP-5-ADJCHANGE: bgp- [29342]
(default) neighbor 100.64.0.202 Up
  
```

- 6. Finalize the configuration of the upper leaf switches (R1UL and R2UL).

Perform this step on the upper leaf switches (R1UL and R2UL) only.

- a. Start the inter-switch ports on the upper leaf switches (R1UL and R2UL).

i. On R1UL:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/4-7, ethernet 1/30-33
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

ii. On R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/1-7, ethernet 1/30-36
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

- b. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- c. Verify the status of the inter-switch ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

i. R1UL# **show interface status**

ii. R2UL# **show interface status**

Examine the output to ensure that the inter-switch ports are connected.

The following example shows the expected output on the X9M rack (R2UL). On the X8M rack (R1UL), ports 01, 02, 03, 34, 35, and 36 are not physically connected.

R2UL# **show interface status**

```
-----
-----
Port          Name          Status      Vlan      Duplex Speed
Type
-----
-----
mgmt0         --            connected   routed    full 1000  --
```



```

-----
-----
Port          Name          Status  Vlan  Duplex
Speed  Type
-----
Eth1/1      RouterPort1   connected  routed  full
100G  QSFP-100G-CR4
Eth1/2      RouterPort2   connected  routed  full
100G  QSFP-100G-CR4
Eth1/3      RouterPort3   connected  routed  full
100G  QSFP-100G-CR4
Eth1/4      RouterPort4   connected  routed  full
100G  QSFP-100G-CR4
Eth1/5      RouterPort5   connected  routed  full
100G  QSFP-100G-CR4
Eth1/6      RouterPort6   connected  routed  full
100G  QSFP-100G-CR4
Eth1/7      RouterPort7   connected  routed  full
100G  QSFP-100G-CR4
Eth1/8      celadm14      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/9      celadm13      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/10     celadm12      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/11     celadm11      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/12     celadm10      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/13     celadm09      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/14     celadm08      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/15     adm08         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/16     adm07         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/17     adm06         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/18     adm05         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/19     adm04         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/20     adm03         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/21     adm02         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/22     adm01         disabled  3888   full
100G  QSFP-100G-CR4
Eth1/23     celadm07      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/24     celadm06      disabled  3888   full
100G  QSFP-100G-CR4
Eth1/25     celadm05      disabled  3888   full
  
```

```

100G    QSFP-100G-CR4
Eth1/26    celadm04          disabled  3888    full 100G
QSFP-100G-CR4
Eth1/27    celadm03          disabled  3888    full 100G
QSFP-100G-CR4
Eth1/28    celadm02          disabled  3888    full 100G
QSFP-100G-CR4
Eth1/29    celadm01          disabled  3888    full 100G
QSFP-100G-CR4
Eth1/30    RouterPort8       connected routed    full 100G
QSFP-100G-CR4
Eth1/31    RouterPort9       connected routed    full 100G
QSFP-100G-CR4
Eth1/32    RouterPort10      connected routed    full 100G
QSFP-100G-CR4
Eth1/33    RouterPort11      connected routed    full 100G
QSFP-100G-CR4
Eth1/34    RouterPort12      connected routed    full 100G
QSFP-100G-CR4
Eth1/35    RouterPort13      connected routed    full 100G
QSFP-100G-CR4
Eth1/36    RouterPort14      connected routed    full 100G
QSFP-100G-CR4
Lo0        Routing loopback i connected routed    auto auto  --
Lo1        VTEP loopback inte connected routed    auto auto  --
Vlan1     --                down    routed    auto auto  --
nve1     --                connected --      auto auto  --
  
```

d. Start the storage server ports on the upper leaf switches (R1UL and R2UL).

i. On R1UL:

```

R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
  
```

ii. Repeat the command sequence on R2UL:

```

R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
  
```

```
R2UL(config)# <Ctrl-Z>
R2UL#
```

- e. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- f. Verify the status of the storage server ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

i. R1UL# **show interface status**

ii. R2UL# **show interface status**

Examine the output to ensure that the storage server ports are `connected`.

The following example shows the expected output on the X9M rack (R2UL). On the X8M rack (R1UL), ports 01, 02, 03, 34, 35, and 36 are not physically connected.

```
R2UL# show interface status
```

```
-----
-----
Port          Name          Status  Vlan  Duplex
Speed  Type
-----
-----
mgmt0        --          connected routed  full
1000        --
-----
-----
Port          Name          Status  Vlan  Duplex
Speed  Type
-----
-----
Eth1/1        RouterPort1   connected routed  full
100G  QSFP-100G-CR4
Eth1/2        RouterPort2   connected routed  full
100G  QSFP-100G-CR4
Eth1/3        RouterPort3   connected routed  full
100G  QSFP-100G-CR4
Eth1/4        RouterPort4   connected routed  full
100G  QSFP-100G-CR4
Eth1/5        RouterPort5   connected routed  full
100G  QSFP-100G-CR4
Eth1/6        RouterPort6   connected routed  full
100G  QSFP-100G-CR4
Eth1/7        RouterPort7   connected routed  full
100G  QSFP-100G-CR4
Eth1/8        celadm14      connected 3888  full
100G  QSFP-100G-CR4
Eth1/9        celadm13      connected 3888  full
100G  QSFP-100G-CR4
```

```
Eth1/10      celadm12      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/11      celadm11      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/12      celadm10      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/13      celadm09      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/14      celadm08      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/15      adm08         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/16      adm07         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/17      adm06         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/18      adm05         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/19      adm04         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/20      adm03         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/21      adm02         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/22      adm01         disabled  3888      full 100G
QSF-P-100G-CR4
Eth1/23      celadm07      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/24      celadm06      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/25      celadm05      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/26      celadm04      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/27      celadm03      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/28      celadm02      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/29      celadm01      connected 3888      full 100G
QSF-P-100G-CR4
Eth1/30      RouterPort8   connected routed      full 100G
QSF-P-100G-CR4
Eth1/31      RouterPort9   connected routed      full 100G
QSF-P-100G-CR4
Eth1/32      RouterPort10  connected routed      full 100G
QSF-P-100G-CR4
Eth1/33      RouterPort11  connected routed      full 100G
QSF-P-100G-CR4
Eth1/34      RouterPort12  connected routed      full 100G
QSF-P-100G-CR4
Eth1/35      RouterPort13  connected routed      full 100G
QSF-P-100G-CR4
Eth1/36      RouterPort14  connected routed      full 100G
QSF-P-100G-CR4
Lo0         Routing loopback i connected routed      auto auto      --
```

```

Lo1          VTEP loopback inte connected routed    auto
auto        --
Vlan1       --                down      routed    auto
auto        --
nve1        --                connected --         auto
auto        --
  
```

g. Start the database server ports on the upper leaf switches (R1UL and R2UL).

i. On R1UL:

```

R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/15-22
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
  
```

ii. Repeat the command sequence on R2UL:

```

R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/15-22
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
  
```

h. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.

i. Verify the status of the database server ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

i. R1UL# **show interface status**

ii. R2UL# **show interface status**

Examine the output to ensure that the database server ports are connected.

The following example shows the expected output on the X9M rack (R2UL). On the X8M rack (R1UL), ports 01, 02, 03, 34, 35, and 36 are not physically connected.

R2UL# **show interface status**

```

-----
-----
Port          Name              Status   Vlan    Duplex Speed
Type
-----
-----
mgmt0         --                connected routed   full 1000  --
-----
-----
Port          Name              Status   Vlan    Duplex Speed
Type
-----
-----
Eth1/1        RouterPort1       connected routed   full 100G
QSFP-100G-CR4
Eth1/2        RouterPort2       connected routed   full 100G
QSFP-100G-CR4
Eth1/3        RouterPort3       connected routed   full 100G
QSFP-100G-CR4
Eth1/4        RouterPort4       connected routed   full 100G
QSFP-100G-CR4
Eth1/5        RouterPort5       connected routed   full 100G
QSFP-100G-CR4
Eth1/6        RouterPort6       connected routed   full 100G
QSFP-100G-CR4
Eth1/7        RouterPort7       connected routed   full 100G
QSFP-100G-CR4
Eth1/8        celadm14          connected 3888    full 100G
QSFP-100G-CR4
Eth1/9        celadm13          connected 3888    full 100G
QSFP-100G-CR4
Eth1/10       celadm12          connected 3888    full 100G
QSFP-100G-CR4
Eth1/11       celadm11          connected 3888    full 100G
QSFP-100G-CR4
Eth1/12       celadm10          connected 3888    full 100G
QSFP-100G-CR4
Eth1/13       celadm09          connected 3888    full 100G
QSFP-100G-CR4
Eth1/14       celadm08          connected 3888    full 100G
QSFP-100G-CR4
Eth1/15       adm08             connected 3888    full 100G
QSFP-100G-CR4
Eth1/16       adm07             connected 3888    full 100G
QSFP-100G-CR4
Eth1/17       adm06             connected 3888    full 100G
QSFP-100G-CR4
Eth1/18       adm05             connected 3888    full 100G
QSFP-100G-CR4
  
```

```

Eth1/19      adm04      connected 3888      full
100G      QSFP-100G-CR4
Eth1/20      adm03      connected 3888      full
100G      QSFP-100G-CR4
Eth1/21      adm02      connected 3888      full
100G      QSFP-100G-CR4
Eth1/22      adm01      connected 3888      full
100G      QSFP-100G-CR4
Eth1/23      celadm07   connected 3888      full
100G      QSFP-100G-CR4
Eth1/24      celadm06   connected 3888      full
100G      QSFP-100G-CR4
Eth1/25      celadm05   connected 3888      full
100G      QSFP-100G-CR4
Eth1/26      celadm04   connected 3888      full
100G      QSFP-100G-CR4
Eth1/27      celadm03   connected 3888      full
100G      QSFP-100G-CR4
Eth1/28      celadm02   connected 3888      full
100G      QSFP-100G-CR4
Eth1/29      celadm01   connected 3888      full
100G      QSFP-100G-CR4
Eth1/30      RouterPort8  connected routed      full
100G      QSFP-100G-CR4
Eth1/31      RouterPort9  connected routed      full
100G      QSFP-100G-CR4
Eth1/32      RouterPort10 connected routed      full
100G      QSFP-100G-CR4
Eth1/33      RouterPort11 connected routed      full
100G      QSFP-100G-CR4
Eth1/34      RouterPort12 connected routed      full
100G      QSFP-100G-CR4
Eth1/35      RouterPort13 connected routed      full
100G      QSFP-100G-CR4
Eth1/36      RouterPort14 connected routed      full
100G      QSFP-100G-CR4
Lo0          Routing loopback i connected routed      auto
auto        --
Lo1          VTEP loopback inte connected routed      auto
auto        --
Vlan1       --          down      routed      auto
auto        --
nve1       --          connected --          auto
auto        --
  
```

j. Verify that nve is up on the leaf switches (R1LL, R1UL, R2LL, and R2UL).

Run the following command on each leaf switch and examine the output:

i. R1LL# **show nve peers**

ii. R1UL# **show nve peers**

iii. R2LL# **show nve peers**

iv. R2UL# **show nve peers**

In the output, you should see three nve peers with State=Up.

For example:

```
R1UL# show nve peers
Interface          Peer-IP State LearnType  Uptime          Router-Mac
-----
nve1               100.64.1.101 Up        CP 00:04:29     n/a
nve1               100.64.1.103 Up        CP 00:07:48     n/a
nve1               100.64.1.104 Up        CP 00:04:10     n/a
```

k. Verify that BGP is up on the upper leaf switches (R1UL and R2UL).

Run the following command on each upper leaf switch and examine the output:

i. R1UL# **show logging log | grep BGP**

ii. R2UL# **show logging log | grep BGP**

In the output, look for two entries with Up in the rightmost column that are associated with different IP addresses.

For example:

```
R1UL# show logging log | grep BGP
2020 Aug 10 13:57:13 R1UL %BGP-5-ADJCHANGE: bgp- [32782] (default)
neighbor 100.64.0.201 Up
2020 Aug 10 13:57:24 R1UL %BGP-5-ADJCHANGE: bgp- [32782] (default)
neighbor 100.64.0.202 Up
```

7. For each rack (R1 and R2), confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

The `verify_roce_cables.py` script uses two input files; one for database servers and storage servers (`nodes.rackN`), and another for switches (`switches.rackN`). In each file, every server or switch must be listed on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

See My Oracle Support document 2587717.1 for download and detailed usage instructions.

Run the `verify_roce_cables.py` script against both of the racks:

```
a. # cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
```

```
b. # cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
```

Check the output of the `verify_roce_cables.py` script against the tables in [Two-Rack Cabling for a System Combining X8M and X9M Racks](#). Also, check that output in the CABLE OK? columns contains the OK status.

The following examples show extracts of the expected command results:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER) LOWER LEAF (rack1sw-rocea0) : CABLE OK? UPPER LEAF
(rack1sw-roceb0) : CABLE OK?
----- : -----
----- : -----
...
```

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
SWITCH PORT (EXPECTED PEER) LOWER LEAF (rack2sw-rocea0) : CABLE OK? UPPER LEAF
(rack2sw-roceb0) : CABLE OK?
----- : -----
----- : -----
...
```

8. Verify the RoCE Network Fabric operation across both interconnected racks by using the `infinicheck` command.

Use the following recommended command sequence to verify the RoCE Network Fabric operation across both racks.

In each command, `hosts.all` contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses from both racks (2 RoCE Network Fabric IP addresses for each database server), and `cells.all` contains a list of RoCE Network Fabric IP addresses for the storage servers from both racks (2 RoCE Network Fabric IP addresses for each storage server).

- a. # cd /opt/oracle.SupportTools/ibdiagtools
./infinicheck -g hosts.all -c cells.all -z
- b. # cd /opt/oracle.SupportTools/ibdiagtools
./infinicheck -g hosts.all -c cells.all -s
- c. # cd /opt/oracle.SupportTools/ibdiagtools
./infinicheck -g hosts.all -c cells.all -b

See step 1.k for most information about each `infinicheck` command.

The following example shows the expected command results for the final command in the sequence:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.all -c cells.all -b

INFINICHECK
      [Network Connectivity, Configuration and
Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
```

```
##### RoCE CONFIGURATION TESTS #####
Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS]... RoCE devices on all DBs and CELLS look good
Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS]... RoCE Policy Routing settings look good
Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS]... RoCE DSCP ToS settings look good
Checking for RoCE PFC settings and DSCP mapping on all DBs and CELLS
[SUCCESS]... RoCE PFC and DSCP settings look good
Checking for RoCE interface MTU settings. Expected value : 2300
[SUCCESS]... RoCE interface MTU settings look good
Verifying switch advertised DSCP on all DBs and CELLS ports ( )
[SUCCESS]... Advertised DSCP settings from RoCE switch looks good
##### CONNECTIVITY TESTS #####
[COMPUTE NODES -> STORAGE CELLS]
(60 seconds approx.)
(Will walk through QoS values: 0-6) [SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
[COMPUTE NODES -> COMPUTE NODES]
...
```

At this point, both racks share the RoCE Network Fabric, and the combined system is ready for further configuration. See [Configuring the New Hardware](#).

- [Two-Rack Cabling for a System Combining X8M and X9M Racks](#)
This section provides the cabling details to connect an X8M rack with an X9M rack, both of which use RoCE Network Fabric.

Related Topics

- [Cabling Two Oracle Exadata Database Machine RoCE Network Fabric Racks With No Downtime \(My Oracle Support Doc ID 2704997.1\)](#)
- [Verify RoCE Cabling on Oracle Exadata Database Machine X8M-2 and X8M-8 Servers \(My Oracle Support Doc ID 2587717.1\)](#)

2.3.2.1.2.1 Two-Rack Cabling for a System Combining X8M and X9M Racks

This section provides the cabling details to connect an X8M rack with an X9M rack, both of which use RoCE Network Fabric.

 **Note:**

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first (X8M) rack is R1, and the second (X9M) rack is R2.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in 2-socket systems (Exadata Database Machine X8M-2 or X9M-2, or Storage Expansion Rack X8M-2 or X9M-2).
 - At rack unit 21 (U21) and rack unit 23 (U23) in 8-socket systems (Exadata Database Machine X8M-8 or X9M-8).
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.

The following illustration shows the cable connections for the spine switches when cabling a two-rack hybrid system with an X8M rack and an X9M rack:



The following tables contain details for all of the RoCE Network Fabric cabling connections in a two-rack hybrid system with an X8M rack and an X9M rack.

Table 2-3 Leaf Switch Connections for the X8M Rack (R1)

Leaf Switch	Connection	Cable Length
R1UL to R1SS	R1UL-P5 to R1SS-P5 R1UL-P7 to R1SS-P7 R1UL-P4 to R1SS-P9 R1UL-P6 to R1SS-P11	3 meters
R1UL to R2SS	R1UL-P31 to R2SS-P5 R1UL-P33 to R2SS-P7 R1UL-P30 to R2SS-P9 R1UL-P32 to R2SS-P11	5 meters
R1LL to R1SS	R1LL-P5 to R1SS-P13 R1LL-P7 to R1SS-P15 R1LL-P4 to R1SS-P17 R1LL-P6 to R1SS-P19	3 meters
R1LL to R2SS	R1LL-P31 to R2SS-P13 R1LL-P33 to R2SS-P15 R1LL-P30 to R2SS-P17 R1LL-P32 to R2SS-P19	5 meters

Table 2-4 Leaf Switch Connections for the X9M Rack (R2)

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8 R2UL-P3 to R1SS-P10 R2UL-P4 to R1SS-P12 R2UL-P5 to R1SS-P14 R2UL-P6 to R1SS-P16 R2UL-P7 to R1SS-P18	5 meters
R2UL to R2SS	R2UL-P30 to R2SS-P6 R2UL-P31 to R2SS-P8 R2UL-P32 to R2SS-P10 R2UL-P33 to R2SS-P12 R2UL-P34 to R2SS-P14 R2UL-P35 to R2SS-P16 R2UL-P36 to R2SS-P18	3 meters

Table 2-4 (Cont.) Leaf Switch Connections for the X9M Rack (R2)

Switch Pair	Connections	Cable Length
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22 R2LL-P3 to R1SS-P24 R2LL-P4 to R1SS-P26 R2LL-P5 to R1SS-P28 R2LL-P6 to R1SS-P30 R2LL-P7 to R1SS-P32	5 meters
R2LL to R2SS	R2LL-P30 to R2SS-P20 R2LL-P31 to R2SS-P22 R2LL-P32 to R2SS-P24 R2LL-P33 to R2SS-P26 R2LL-P34 to R2SS-P28 R2LL-P35 to R2SS-P30 R2LL-P36 to R2SS-P32	3 meters

2.3.2.1.3 Extending an X8M Rack with No Down Time by Adding Another X8M Rack

 **WARNING:**

Take time to read and understand this procedure before implementation. Pay careful attention to the instructions that surround the command examples. A system outage may occur if the procedure is not applied correctly.

 **Note:**

This procedure assumes that the RoCE Network Fabric switches on the X8M racks contain the golden configuration settings from Oracle Exadata System Software 20.1.0 or later. Otherwise, before using this procedure, you must update the Oracle Exadata System Software and update the golden configuration settings on the RoCE Network Fabric switches. Downtime is required to update the golden configuration settings on the RoCE Network Fabric switches.

 **Note:**

For additional background information, see [Understanding Multi-Rack Cabling for X8M Racks](#).

Use this procedure to extend a typical X8M rack without down-time by cabling it together with a second X8M rack. The primary rack (designated R1) and all of the systems it supports remain online throughout the procedure. At the beginning of the procedure, the additional rack (designated R2) is shut down.

The following is an outline of the procedure:

- Preparation (steps 1 and 2)

In this phase, you prepare the racks, switches, and cables. Also, you install and cable the spine switches in both racks.

- Configuration and Physical Cabling

In this phase, you reconfigure the leaf switches and finalize the cabling to the spine switches. These tasks are carefully orchestrated to avoid downtime on the primary system, as follows:

- Partially configure the lower leaf switches (step 3)

In this step, you reconfigure the switch ports on the lower leaf switches. There is no physical cabling performed in this step.

- Partially configure the upper leaf switches (step 4)

In this step, you reconfigure the switch ports on the upper leaf switches, remove the inter-switch cables that connect the leaf switches in both racks and connect the cables between the upper leaf switches and the spine switches.

- Finalize the lower leaf switches (step 5)

In this step, you finalize the switch port configuration on the lower leaf switches. You also complete the physical cabling by connecting the cables between the lower leaf switches and the spine switches.

- Finalize the upper leaf switches (step 6)

In this step, you finalize the switch port configuration on the upper leaf switches.

- Validation and Testing (steps 7 and 8)

In this phase, you validate and test the RoCE Network Fabric across both of the interconnect racks.

After completing the procedure, both racks share the RoCE Network Fabric, and the combined system is ready for further configuration. For example, you can extend existing disk groups and Oracle RAC databases to consume resources across both racks.

 **Note:**

- This procedure applies only to typical rack configurations that initially have leaf switches with the following specifications:
 - The inter-switch ports are ports 4 to 7, and ports 30 to 33.
 - The storage server ports are ports 8 to 14, and ports 23 to 29.
 - The database server ports are ports 15 to 22.

For other rack configurations (for example, X8M-8 systems with three database servers and 11 storage servers) a different procedure and different RoCE Network Fabric switch configuration files are required. Contact Oracle for further guidance.

- The procedure uses the following naming abbreviations and conventions:
 - The abbreviation for the existing rack is R1, and the new rack is R2.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies a spine switch.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the existing rack (R1).
- Most operations must be executed in multiple locations. For example, step 1.h instructs you to update the firmware on all the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL). Pay attention to the instructions and keep track of your actions.

 **Tip:**

When a step must be performed on multiple switches, the instruction contains a list of the applicable switches. For example, (R1LL, R1UL, R2LL, and R2UL). You can use this list as a checklist to keep track of your actions.

- Perform operations sequentially, and complete every operation before proceeding. For example, run the entire command sequence at 3.a.i as one operation and complete it before proceeding.
- All of commands that are run on a RoCE Network Fabric switch must be run while connected to the switch management interface as the switch administrator.

1. Prepare the systems.

- a. Position the new rack (R2) so that it is physically near the existing rack (R1). The RDMA Network Fabric cables must be able to reach the switches in each rack.

For the required cross-rack cabling information, see [Two-Rack Cabling for X8M Racks](#).

- b. Power on all of the servers and network switches in the new rack (R2).
This includes the database servers, storage servers, RoCE Network Fabric leaf switches, and the Management Network Switch.
- c. Prepare the RoCE Network Fabric cables that you will use to interconnect the racks.
Label both ends of every cable.
For the required cross-rack cabling information, see [Two-Rack Cabling for X8M Racks](#).
- d. Connect the new rack (R2) to your existing management network.
Ensure that there are no IP address conflicts across the racks and that you can access the management interfaces on the RoCE Network Fabric switches.
- e. Ensure that you have a backup of the current switch configuration for each RoCE Network Fabric switch (R1LL, R1UL, R2LL, and R2UL).
See [Backing Up Settings on the RoCE Network Fabric Switch](#) in *Oracle Exadata Database Machine Maintenance Guide*.
- f. Download the required RoCE Network Fabric switch configuration files.
This procedure requires specific RoCE Network Fabric switch configuration files, which you must download from [My Oracle Support document 2704997.1](#).

 **WARNING:**

You must use different switch configuration files depending on whether your system uses Exadata Secure RDMA Fabric Isolation. Ensure that you download the correct archive that matches your system configuration.

For system configurations without Secure Fabric, download `online_multi-rack.zip`. For system configurations with Secure Fabric, download `online_SF_enabled_multi-rack.zip`.

Download and extract the archive containing the required RoCE Network Fabric switch configuration files. Place the files on a server with access to the management interfaces on the RoCE Network Fabric switches.

- g. Copy the required RoCE Network Fabric switch configuration files to the leaf switches on both racks.

You can use the following commands to copy the required configuration files to all of the RoCE Network Fabric switches on a system without Secure Fabric enabled:

- i.

```
# scp roce_multi_online_step3_R1_LL.cfg admin@R1LL_IP:/
```
- ii.

```
# scp roce_multi_online_step3_R2_LL.cfg admin@R2LL_IP:/
```
- iii.

```
# scp roce_multi_online_step4_R1_UL.cfg admin@R1UL_IP:/
```



```
iv. # scp roce_multi_online_step4_R2_UL.cfg admin@R2UL_IP:/
```

```
v. # scp roce_multi_online_step5.cfg admin@R1LL_IP:/
```

```
vi. # scp roce_multi_online_step5.cfg admin@R2LL_IP:/
```

On a system with Secure Fabric enabled, you can use the following commands:

```
i. # scp roce_SF_multi_online_step3_R1_LL.cfg admin@R1LL_IP:/
```

```
ii. # scp roce_SF_multi_online_step3_R2_LL.cfg admin@R2LL_IP:/
```

```
iii. # scp roce_SF_multi_online_step4_R1_UL.cfg admin@R1UL_IP:/
```

```
iv. # scp roce_SF_multi_online_step4_R2_UL.cfg admin@R2UL_IP:/
```

```
v. # scp roce_SF_multi_online_step5.cfg admin@R1LL_IP:/
```

```
vi. # scp roce_SF_multi_online_step5.cfg admin@R2LL_IP:/
```

In the above commands, substitute the appropriate IP address or host name where applicable. For example, in place of *R1LL_IP*, substitute the management IP address or host name for the lower leaf switch (LL) on the existing rack (R1).

 **Note:**

The command examples in the rest of this procedure use the configuration files for a system configuration without Secure Fabric enabled. If required, adjust the commands to use the Secure Fabric-enabled switch configuration files.

- h. Update the firmware to the latest available release on all of the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL).

See Updating RoCE Network Fabric Switch Firmware in *Oracle Exadata Database Machine Maintenance Guide*.

- i. Examine the RoCE Network Fabric leaf switches (R1LL, R1UL, R2LL, and R2UL) and confirm the port categories for the cabled ports.

Run the `show interface status` command on every RoCE Network Fabric leaf switch:

```
i. R1LL# show interface status
```

```
ii. R1UL# show interface status
```

```
iii. R2LL# show interface status
```

iv. R2U1# show interface status

Examine the output and confirm the port categories as follows:

- Confirm that the inter-switch ports are ports 4 to 7, and ports 30 to 33.
- Confirm that the storage server ports are ports 8 to 14, and ports 23 to 29.
- Confirm that the database server ports are ports 15 to 22.

For example:

R1LL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
Eth1/1        --            xcvrAbsen  1          auto    auto    --
Eth1/2        --            xcvrAbsen  1          auto    auto    --
Eth1/3        --            xcvrAbsen  1          auto    auto    --
Eth1/4        ISL1          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/5        ISL2          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/6        ISL3          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/7        ISL4          connected   trunk     full    100G
QSFP-100G-CR4
Eth1/8        celadm14     connected   3888     full    100G
QSFP-100G-CR4
Eth1/9        celadm13     connected   3888     full    100G
QSFP-100G-CR4
Eth1/10       celadm12     connected   3888     full    100G
QSFP-100G-CR4
Eth1/11       celadm11     connected   3888     full    100G
QSFP-100G-CR4
Eth1/12       celadm10     connected   3888     full    100G
QSFP-100G-CR4
Eth1/13       celadm09     connected   3888     full    100G
QSFP-100G-CR4
Eth1/14       celadm08     connected   3888     full    100G
QSFP-100G-CR4
Eth1/15       adm08        connected   3888     full    100G
QSFP-100G-CR4
Eth1/16       adm07        connected   3888     full    100G

```

```

QSF1P-100G-CR4
Eth1/17      adm06      connected 3888      full
100G QSF1P-100G-CR4
Eth1/18      adm05      connected 3888      full
100G QSF1P-100G-CR4
Eth1/19      adm04      connected 3888      full
100G QSF1P-100G-CR4
Eth1/20      adm03      connected 3888      full
100G QSF1P-100G-CR4
Eth1/21      adm02      connected 3888      full
100G QSF1P-100G-CR4
Eth1/22      adm01      connected 3888      full
100G QSF1P-100G-CR4
Eth1/23      celadm07   connected 3888      full
100G QSF1P-100G-CR4
Eth1/24      celadm06   connected 3888      full
100G QSF1P-100G-CR4
Eth1/25      celadm05   connected 3888      full
100G QSF1P-100G-CR4
Eth1/26      celadm04   connected 3888      full
100G QSF1P-100G-CR4
Eth1/27      celadm03   connected 3888      full
100G QSF1P-100G-CR4
Eth1/28      celadm02   connected 3888      full
100G QSF1P-100G-CR4
Eth1/29      celadm01   connected 3888      full
100G QSF1P-100G-CR4
Eth1/30      ISL5       connected trunk    full
100G QSF1P-100G-CR4
Eth1/31      ISL6       connected trunk    full
100G QSF1P-100G-CR4
Eth1/32      ISL7       connected trunk    full
100G QSF1P-100G-CR4
Eth1/33      ISL8       connected trunk    full
100G QSF1P-100G-CR4
Eth1/34      --         xcvrAbsen 1       auto
auto --
Eth1/35      --         xcvrAbsen 1       auto
auto --
Eth1/36      --         xcvrAbsen 1       auto
auto --
Po100      --         connected trunk    full
100G --
Lo0        Routing loopback i connected routed    auto
auto --
Lo1        VTEP loopback inte connected routed    auto
auto --
Vlan1      --         down      routed    auto
auto --
nve1      --         connected --      auto
auto --

```

- j. For each rack (R1 and R2), confirm the RoCE Network Fabric cabling by running the `verify_roce_cables.py` script.

The `verify_roce_cables.py` script uses two input files; one for database servers and storage servers (`nodes.rackN`), and another for switches (`switches.rackN`). In each file, every server or switch must be listed on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

See My Oracle Support document 2587717.1 for download and detailed usage instructions.

Run the `verify_roce_cables.py` script against both of the racks:

- i. `# cd /opt/oracle.SupportTools/ibdiagtools`
`# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1`
- ii. `# cd /opt/oracle.SupportTools/ibdiagtools`
`# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2`

Check that output in the `CABLE OK?` columns contains the `OK` status.

The following example shows the expected command results:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER)                LOWER LEAF (rack1sw-
rocea0) : CABLE OK?                          UPPER LEAF (rack1sw-roceb0) : CABLE
OK?
-----
----- : -----
----- : -----
      Eth1/4 (ISL peer switch) : rack1sw-rocea0
Ethernet1/4 : OK                rack1sw-roceb0 Ethernet1/4 : OK
      Eth1/5 (ISL peer switch) : rack1sw-rocea0
Ethernet1/5 : OK                rack1sw-roceb0 Ethernet1/5 : OK
      Eth1/6 (ISL peer switch) : rack1sw-rocea0
Ethernet1/6 : OK                rack1sw-roceb0 Ethernet1/6 : OK
      Eth1/7 (ISL peer switch) : rack1sw-rocea0
Ethernet1/7 : OK                rack1sw-roceb0 Ethernet1/7 : OK
      Eth1/8 (RU39)            : rack1celadm14
port-1 : OK                    rack1celadm14 port-2 : OK
      Eth1/9 (RU37)            : rack1celadm13
port-1 : OK                    rack1celadm13 port-2 : OK
      Eth1/10 (RU35)           : rack1celadm12
port-1 : OK                    rack1celadm12 port-2 : OK
      Eth1/11 (RU33)           : rack1celadm11
port-1 : OK                    rack1celadm11 port-2 : OK
      Eth1/12 (RU31)           : rack1celadm10
port-1 : OK                    rack1celadm10 port-2 : OK
      Eth1/13 (RU29)           : rack1celadm09
port-1 : OK                    rack1celadm09 port-2 : OK
      Eth1/14 (RU27)           : rack1celadm08
port-1 : OK                    rack1celadm08 port-2 : OK
      Eth1/15 (RU26)           : rack1adm08
port-1 : OK                    rack1adm08 port-2 : OK
      Eth1/16 (RU25)           : rack1adm07
port-1 : OK                    rack1adm07 port-2 : OK
      Eth1/17 (RU24)           : rack1adm06
port-1 : OK                    rack1adm06 port-2 : OK
```

```

    Eth1/18 (RU23) : rack1adm5
port-1 : OK rack1adm5 port-2 : OK
    Eth1/19 (RU19) : rack1adm4
port-1 : OK rack1adm4 port-2 : OK
    Eth1/20 (RU18) : rack1adm3
port-1 : OK rack1adm3 port-2 : OK
    Eth1/21 (RU17) : rack1adm2
port-1 : OK rack1adm2 port-2 : OK
    Eth1/22 (RU16) : rack1adm1
port-1 : OK rack1adm1 port-2 : OK
    Eth1/23 (RU14) : rack1celadm7
port-1 : OK rack1celadm7 port-2 : OK
    Eth1/24 (RU12) : rack1celadm6
port-1 : OK rack1celadm6 port-2 : OK
    Eth1/25 (RU10) : rack1celadm5
port-1 : OK rack1celadm5 port-2 : OK
    Eth1/26 (RU08) : rack1celadm4
port-1 : OK rack1celadm4 port-2 : OK
    Eth1/27 (RU06) : rack1celadm3
port-1 : OK rack1celadm3 port-2 : OK
    Eth1/28 (RU04) : rack1celadm2
port-1 : OK rack1celadm2 port-2 : OK
    Eth1/29 (RU02) : rack1celadm1
port-1 : OK rack1celadm1 port-2 : OK
    Eth1/30 (ISL peer switch) : rack1sw-rocea0
Ethernet1/30 : OK rack1sw-roceb0
Ethernet1/30 : OK
    Eth1/31 (ISL peer switch) : rack1sw-rocea0
Ethernet1/31 : OK rack1sw-roceb0
Ethernet1/31 : OK
    Eth1/32 (ISL peer switch) : rack1sw-rocea0
Ethernet1/32 : OK rack1sw-roceb0
Ethernet1/32 : OK
    Eth1/33 (ISL peer switch) : rack1sw-rocea0
Ethernet1/33 : OK rack1sw-roceb0
Ethernet1/33 : OK
  
```

- k. For each rack (R1 and R2), verify the RoCE Network Fabric operation by using the `infinicheck` command.
- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command.
 - Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric.
 - Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks.

In each command, the `hosts` input file (`hosts.rack1` and `hosts.rack2`) contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses (2 RoCE Network Fabric IP addresses for each database server), and the `cells` input file (`cells.rack1` and `cells.rack2`)

contains a list of RoCE Network Fabric IP addresses for the storage servers (2 RoCE Network Fabric IP addresses for each storage server).

i. Use the following recommended command sequence on the existing rack (R1):

```
i. # cd /opt/oracle.SupportTools/ibdiagtools
   # ./infinicheck -g hosts.rack1 -c cells.rack1 -z

ii. # cd /opt/oracle.SupportTools/ibdiagtools
    # ./infinicheck -g hosts.rack1 -c cells.rack1 -s

iii. # cd /opt/oracle.SupportTools/ibdiagtools
     # ./infinicheck -g hosts.rack1 -c cells.rack1 -b
```

ii. Use the following recommended command sequence on the new rack (R2):

```
i. # cd /opt/oracle.SupportTools/ibdiagtools
   # ./infinicheck -g hosts.rack2 -c cells.rack2 -z

ii. # cd /opt/oracle.SupportTools/ibdiagtools
    # ./infinicheck -g hosts.rack2 -c cells.rack2 -s

iii. # cd /opt/oracle.SupportTools/ibdiagtools
     # ./infinicheck -g hosts.rack2 -c cells.rack2 -b
```

The following example shows the expected command results for the final command in the sequence:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.rackN -c cells.rackN -b

INFINICHECK
      [Network Connectivity, Configuration and Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS]... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS]... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS]... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs and
CELLS
[SUCCESS]... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value : 2300
[SUCCESS]... RoCE interface MTU settings look good
      Verifying switch advertised DSCP on all DBs and CELLS ports ( )
[SUCCESS]... Advertised DSCP settings from RoCE switch looks good
      ##### CONNECTIVITY TESTS #####
      [COMPUTE NODES -> STORAGE CELLS]
```

```
(60 seconds approx.)
(Will walk through QoS values: 0-6)
[SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
[COMPUTE NODES -> COMPUTE NODES]
...
```

2. Install the spine switches (R1SS and R2SS).
 - a. Physically install and power up the spine switches in the existing rack (R1SS) and the new rack (R2SS).
 - i. Physically install each spine switch in RU1.
 - ii. For each spine switch, ensure that the management Ethernet interface is connected to the management network and then supply power.
 - iii. On each spine switch, perform the initial configuration steps outlined in *Configuring the Cisco Nexus C9336C-FX2 Switch*. Skip the step for applying the golden configuration settings as you will do this later.
 - iv. For each spine switch, perform a ping test to the management Ethernet interface to ensure that the switch is online and accessible.

- b. Apply the golden configuration settings to the new spine switches.

See *Applying Golden Configuration Settings on RoCE Network Fabric Switches* in *Oracle Exadata Database Machine Maintenance Guide*.

You can use the instance of `patchmgr` that you previously used to update the firmware on the leaf switches (in step 1.h).

Use a switch list file (`spines.lst`) to apply the golden configuration settings to both spine switches using one `patchmgr` command:

```
# cat spines.lst
R1SS_IP:m spine.201
R2SS_IP:m spine.202

# ./patchmgr --roceswitches spines.lst --apply-config -
log_dir /tmp/spinelogs
```

 **Note:**

In the switch list file, `R1SS_IP` is the management IP address or host name for the spine switch on the existing rack (R1SS) and `R2SS_IP` is the management IP address or host name for the spine switch on the new rack (R2SS).

- c. Upgrade the firmware on the spine switches.

See *Updating RoCE Network Fabric Switch Firmware* in *Oracle Exadata Database Machine Maintenance Guide*.

You can use the instance of `patchmgr` that you used in the previous step.

Use a switch list file (`spines.lst`) to perform the firmware upgrade on both spine switches using one `patchmgr` command:

```
# cat spines.lst
R1SS_IP:mspine.201
R2SS_IP:mspine.202

# ./patchmgr --roceswitches spines.lst --upgrade -log_dir /tmp/
spinelogs
```

 **Note:**

In the switch list file, `R1SS_IP` is the management IP address or host name for the spine switch on the existing rack (R1SS) and `R2SS_IP` is the management IP address or host name for the spine switch on the new rack (R2SS).

- d. Connect the RoCE Network Fabric cables to the spine switches (R1SS and R2SS).

 **WARNING:**

**At this stage, only connect the cables to the spine switches.
DO NOT CONNECT ANY OF THE CABLES TO THE LEAF SWITCHES.**

Use the cabled that you prepared earlier (in step 1.c).

For the required cross-rack cabling information, see [Two-Rack Cabling for X8M Racks](#).

3. Perform the first round of configuration on the lower leaf switches (R1LL and R2LL).
Perform this step on the lower leaf switches (R1LL and R2LL) only.

 **Note:**

During this step, the lower leaf switch ports are shut down. While the R1LL ports are down, R1UL exclusively supports the RoCE Network Fabric. During this time, there is no redundancy in the RoCE Network Fabric, and availability cannot be maintained if R1UL goes down.

- a. Shut down the switch ports on the lower leaf switches (R1LL and R2LL).
 - i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/4-33
R1LL(config-if-range)# shut
R1LL(config-if-range)# exit
```



```
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/4-33
R2LL(config-if-range)# shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- b. Reconfigure the lower leaf switch ports (R1LL and R2LL) .

For each switch, you must use the correct corresponding switch configuration file, which you earlier copied to the switch (in step 1.g).

- i. On R1LL, the switch configuration file name must end with `step3_R1_LL.cfg`:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# run-script bootflash:///
roce_multi_online_step3_R1_LL.cfg | grep 'none'
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. On R2LL, the switch configuration file name must end with `step3_R2_LL.cfg`:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# run-script bootflash:///
roce_multi_online_step3_R2_LL.cfg | grep 'none'
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- c. Start the inter-switch ports on the lower leaf switches (R1LL and R2LL) .

- i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/4-7, ethernet 1/30-33
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

- ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/4-7, ethernet 1/30-33
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- d. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- e. Verify the status of the inter-switch ports on the lower leaf switches (R1LL and R2LL) .

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**
- ii. R2LL# **show interface status**

Examine the output to ensure that the inter-switch ports are connected.

For example:

```
R1LL# show interface status
```

```
-----
-----
```

Port Speed	Name Type	Status	Vlan	Duplex
mgmt0 1000	-- --	connected	routed	full

Port Speed	Name Type	Status	Vlan	Duplex
Eth1/1 auto	-- --	xcvrAbsen	1	auto
Eth1/2 auto	-- --	xcvrAbsen	1	auto
Eth1/3 auto	-- --	xcvrAbsen	1	auto
Eth1/4 100G	ISL1 QSFP-100G-CR4	connected	trunk	full
Eth1/5 100G	ISL2 QSFP-100G-CR4	connected	trunk	full
Eth1/6 100G	ISL3 QSFP-100G-CR4	connected	trunk	full
Eth1/7 100G	ISL4 QSFP-100G-CR4	connected	trunk	full
Eth1/8 100G	celadm14 QSFP-100G-CR4	disabled	3888	full
Eth1/9 100G	celadm13 QSFP-100G-CR4	disabled	3888	full
Eth1/10 100G	celadm12 QSFP-100G-CR4	disabled	3888	full
Eth1/11 100G	celadm11 QSFP-100G-CR4	disabled	3888	full
Eth1/12 100G	celadm10 QSFP-100G-CR4	disabled	3888	full
Eth1/13 100G	celadm09 QSFP-100G-CR4	disabled	3888	full
Eth1/14 100G	celadm08 QSFP-100G-CR4	disabled	3888	full
Eth1/15 100G	adm08 QSFP-100G-CR4	disabled	3888	full
Eth1/16 100G	adm07 QSFP-100G-CR4	disabled	3888	full
Eth1/17 100G	adm06 QSFP-100G-CR4	disabled	3888	full
Eth1/18 100G	adm05 QSFP-100G-CR4	disabled	3888	full
Eth1/19 100G	adm04 QSFP-100G-CR4	disabled	3888	full
Eth1/20 100G	adm03 QSFP-100G-CR4	disabled	3888	full
Eth1/21 100G	adm02 QSFP-100G-CR4	disabled	3888	full

Eth1/22	adm01	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/23	celadm07	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/24	celadm06	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/25	celadm05	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/26	celadm04	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/27	celadm03	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/28	celadm02	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/29	celadm01	disabled	3888	full	100G		
QSF100G-CR4							
Eth1/30	ISL5	connected	trunk	full	100G		
QSF100G-CR4							
Eth1/31	ISL6	connected	trunk	full	100G		
QSF100G-CR4							
Eth1/32	ISL7	connected	trunk	full	100G		
QSF100G-CR4							
Eth1/33	ISL8	connected	trunk	full	100G		
QSF100G-CR4							
Eth1/34	--	xcvrAbsen	1	auto	auto	--	
Eth1/35	--	xcvrAbsen	1	auto	auto	--	
Eth1/36	--	xcvrAbsen	1	auto	auto	--	
Po100	--	connected	trunk	full	100G	--	
Lo0	Routing loopback	i	connected	routed	auto	auto	--
Lo1	VTEP loopback	inte	connected	routed	auto	auto	--
Vlan1	--	down	routed	auto	auto	--	
nve1	--	connected	--	auto	auto	--	

f. Start the storage server ports on the lower leaf switches (R1LL and R2LL) .

i. On R1LL:

```
R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
```

ii. Repeat the command sequence on R2LL:

```
R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R2LL(config-if-range)# no shut
```

```
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
```

- g.** Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- h.** Verify the status of the storage server ports on the lower leaf switches (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

i. R1LL# `show interface status`

ii. R2LL# `show interface status`

Examine the output to ensure that the storage server ports are connected.

For example:

R1LL# `show interface status`

```
-----
-----
Port          Name          Status      Vlan      Duplex
Speed      Type
-----
-----
mgmt0          --            connected   routed    full
1000          --
-----
-----
Port          Name          Status      Vlan      Duplex
Speed      Type
-----
-----
Eth1/1          --            xcvrAbsen  1          auto
auto          --
Eth1/2          --            xcvrAbsen  1          auto
auto          --
Eth1/3          --            xcvrAbsen  1          auto
auto          --
Eth1/4          ISL1          connected   trunk     full
100G          QSFP-100G-CR4
Eth1/5          ISL2          connected   trunk     full
100G          QSFP-100G-CR4
Eth1/6          ISL3          connected   trunk     full
100G          QSFP-100G-CR4
Eth1/7          ISL4          connected   trunk     full
100G          QSFP-100G-CR4
Eth1/8          celadm14      connected   3888      full
```

100G	QSF	100G-CR4						
Eth1/9	celadm13		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/10	celadm12		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/11	celadm11		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/12	celadm10		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/13	celadm09		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/14	celadm08		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/15	adm08		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/16	adm07		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/17	adm06		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/18	adm05		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/19	adm04		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/20	adm03		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/21	adm02		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/22	adm01		disabled	3888	full	100G		
QSF	100G-CR4							
Eth1/23	celadm07		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/24	celadm06		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/25	celadm05		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/26	celadm04		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/27	celadm03		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/28	celadm02		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/29	celadm01		connected	3888	full	100G		
QSF	100G-CR4							
Eth1/30	ISL5		connected	trunk	full	100G		
QSF	100G-CR4							
Eth1/31	ISL6		connected	trunk	full	100G		
QSF	100G-CR4							
Eth1/32	ISL7		connected	trunk	full	100G		
QSF	100G-CR4							
Eth1/33	ISL8		connected	trunk	full	100G		
QSF	100G-CR4							
Eth1/34	--		xcvrAbsen	1	auto	auto	--	
Eth1/35	--		xcvrAbsen	1	auto	auto	--	
Eth1/36	--		xcvrAbsen	1	auto	auto	--	
Po100	--		connected	trunk	full	100G	--	

```

Lo0          Routing loopback i connected routed    auto
auto        --
Lo1          VTEP loopback inte connected routed    auto
auto        --
Vlan1       --                down      routed    auto
auto        --
nve1        --                connected --         auto
auto        --
  
```

- i. Start the database server ports on the lower leaf switches (R1LL and R2LL).

- i. On R1LL:

```

R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# interface ethernet 1/15-22
R1LL(config-if-range)# no shut
R1LL(config-if-range)# exit
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
  
```

- ii. Repeat the command sequence on R2LL:

```

R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# interface ethernet 1/15-22
R2LL(config-if-range)# no shut
R2LL(config-if-range)# exit
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
  
```

- j. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- k. Verify the status of the database server ports on the lower leaf switches (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# **show interface status**

- ii. R2LL# **show interface status**

Examine the output to ensure that the database server ports are connected.

For example:

R1LL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
Eth1/1        --            xcvrAbsen  1         auto    auto    --
Eth1/2        --            xcvrAbsen  1         auto    auto    --
Eth1/3        --            xcvrAbsen  1         auto    auto    --
Eth1/4        ISL1          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/5        ISL2          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/6        ISL3          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/7        ISL4          connected   trunk     full    100G
QSF1P-100G-CR4
Eth1/8        celadm14     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/9        celadm13     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/10       celadm12     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/11       celadm11     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/12       celadm10     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/13       celadm09     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/14       celadm08     connected   3888     full    100G
QSF1P-100G-CR4
Eth1/15       adm08        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/16       adm07        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/17       adm06        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/18       adm05        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/19       adm04        connected   3888     full    100G
QSF1P-100G-CR4
Eth1/20       adm03        connected   3888     full    100G
QSF1P-100G-CR4

```



```

Eth1/21      adm02      connected 3888      full
100G      QSFP-100G-CR4
Eth1/22      adm01      connected 3888      full
100G      QSFP-100G-CR4
Eth1/23      celadm07   connected 3888      full
100G      QSFP-100G-CR4
Eth1/24      celadm06   connected 3888      full
100G      QSFP-100G-CR4
Eth1/25      celadm05   connected 3888      full
100G      QSFP-100G-CR4
Eth1/26      celadm04   connected 3888      full
100G      QSFP-100G-CR4
Eth1/27      celadm03   connected 3888      full
100G      QSFP-100G-CR4
Eth1/28      celadm02   connected 3888      full
100G      QSFP-100G-CR4
Eth1/29      celadm01   connected 3888      full
100G      QSFP-100G-CR4
Eth1/30      ISL5       connected trunk    full
100G      QSFP-100G-CR4
Eth1/31      ISL6       connected trunk    full
100G      QSFP-100G-CR4
Eth1/32      ISL7       connected trunk    full
100G      QSFP-100G-CR4
Eth1/33      ISL8       connected trunk    full
100G      QSFP-100G-CR4
Eth1/34      --         xcvrAbsen 1       auto
auto      --
Eth1/35      --         xcvrAbsen 1       auto
auto      --
Eth1/36      --         xcvrAbsen 1       auto
auto      --
Po100      --         connected trunk    full
100G      --
Lo0        Routing loopback i connected routed    auto
auto      --
Lo1        VTEP loopback inte connected routed    auto
auto      --
Vlan1      --         down      routed    auto
auto      --
nve1      --         connected --      auto
auto      --
  
```

 **Note:**

Before proceeding, ensure that you have completed all of the actions in step 3 on both lower leaf switches (R1LL and R2LL). If not, then ensure that you go back and perform the missing actions.

4. Perform the first round of configuration on the upper leaf switches (R1UL and R2UL).

Perform this step on the upper leaf switches (R1UL and R2UL) only.

 **Note:**

At the start of this step, the upper leaf switch ports are shut down. While the R1UL ports are down, R1LL exclusively supports the RoCE Network Fabric on the existing rack. During this time, there is no redundancy in the RoCE Network Fabric, and availability cannot be maintained if R1LL goes down.

- a. Shut down the upper leaf switch ports (R1UL and R2UL).

- i. On R1UL:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/4-33
R1UL(config-if-range)# shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. Repeat the command sequence on R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/4-33
R2UL(config-if-range)# shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

- b. On both racks, remove the inter-switch links between the leaf switches (R1LL to R1UL, and R2LL to R2UL).

On every leaf switch, remove the cables for the inter-switch links:

- i. On R1LL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
 - ii. On R1UL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
 - iii. On R2LL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
 - iv. On R2UL, disconnect the inter-switch links from ports 04, 05, 06, 07, 30, 31, 32, and 33.
- c. On both racks, cable the upper leaf switch to both of the spine switches (R1UL and R2UL to R1SS and R2SS).

Connect the cables from the spine switches that you prepared earlier (in step 2.d).

Cable the switches as described in [Two-Rack Cabling for X8M Racks](#):

- i. On R1UL, cable ports 04, 05, 06, 07, 30, 31, 32, and 33 to R1SS and R2SS.
 - ii. On R2UL, cable ports 04, 05, 06, 07, 30, 31, 32, and 33 to R1SS and R2SS.
- d. Reconfigure the upper leaf switch ports (R1UL and R2UL).

For each switch, you must use the correct corresponding switch configuration file, which you earlier copied to the switch (in step 1.g):

- i. On R1UL, the switch configuration file name must end with `step4_R1_UL.cfg`:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# run-script bootflash:///
roce_multi_online_step4_R1_UL.cfg | grep 'none'
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. On R2UL, the switch configuration file name must end with `step4_R2_UL.cfg`:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# run-script bootflash:///
roce_multi_online_step4_R2_UL.cfg | grep 'none'
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- e. Check the status of the RoCE Network Fabric ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

- i. R1UL# **show interface status**

ii. R2UL# **show interface status**

Examine the output to ensure that all of the cabled ports are disabled.

For example:

R1UL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
mgmt0         --            connected  routed   full    1000
--
  
```

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
...
Eth1/4        RouterPort1   disabled    routed    full    100G
QSFP-100G-CR4
Eth1/5        RouterPort2   disabled    routed    full    100G
QSFP-100G-CR4
Eth1/6        RouterPort3   disabled    routed    full    100G
QSFP-100G-CR4
Eth1/7        RouterPort4   disabled    routed    full    100G
QSFP-100G-CR4
Eth1/8        celadm14      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/9        celadm13      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/10       celadm12      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/11       celadm11      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/12       celadm10      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/13       celadm09      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/14       celadm08      disabled    3888     full    100G
QSFP-100G-CR4
Eth1/15       adm08         disabled    3888     full    100G
QSFP-100G-CR4
Eth1/16       adm07         disabled    3888     full    100G
QSFP-100G-CR4
Eth1/17       adm06         disabled    3888     full    100G
QSFP-100G-CR4
Eth1/18       adm05         disabled    3888     full    100G
  
```

```
QSFP-100G-CR4
...
```

 **Note:**

Before proceeding, ensure that you have completed all of the actions to this point in step 4 on both upper leaf switches (R1UL and R2UL). If not, then ensure that you go back and perform the missing actions.

f. Verify the configuration of the upper leaf switches.

You can use the instance of `patchmgr` that you previously used to update the switch firmware (in step 1.h).

Use a switch list file (`ul.lst`) to check both upper leaf switches using one `patchmgr` command:

```
# cat ul.lst
R1UL_IP:mleaf.102
R2UL_IP:mleaf.104
```

On a system with Secure Fabric enabled, use the `msfleaf` tag in the switch list file:

```
# cat ul.lst
R1UL_IP:msfleaf.102
R2UL_IP:msfleaf.104
```

The following shows the recommended command and an example of the expected results:

```
# ./patchmgr --roceswitches ul.lst --verify-config -
log_dir /tmp/log
2020-08-10 13:40:09 -0700      :Working: Initiating config
verification... Expect up to 6 minutes for each switch
Mon Aug 10 13:40:13 PDT 2020 1 of 4 :Verifying config on
switch ...
...
Mon Aug 10 13:40:32 PDT 2020:      [INFO      ] Config matches
template: ...
Mon Aug 10 13:40:32 PDT 2020:      [SUCCESS   ] Config validation
successful!
2020-08-10 13:40:32 -0700
  Config check on RoCE switch(es)
2020-08-10 13:40:32 -0700
  Completed run of command: ./patchmgr --roceswitches ul.lst --
verify-config -log_dir /tmp/log
2020-08-10 13:40:32 -0700      :INFO      : config attempted on
nodes in file ul.lst: [R1UL_IP R2UL_IP]
2020-08-10 13:40:32 -0700      :INFO      : For details, check
the following files in /tmp/log:
2020-08-10 13:40:32 -0700      :INFO      : -
```

```

updateRoceSwitch.log
2020-08-10 13:40:32 -0700      :INFO      : - updateRoceSwitch.trc
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.stdout
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.stderr
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.log
2020-08-10 13:40:32 -0700      :INFO      : - patchmgr.trc
2020-08-10 13:40:32 -0700      :INFO      : Exit status:0
2020-08-10 13:40:32 -0700      :INFO      : Exiting.
  
```

In the command output, verify that the switch configuration is good for both upper leaf switches. You can ignore messages about the ports that are down.

5. Finalize the configuration of the lower leaf switches (R1LL and R2LL).

Perform this step on the lower leaf switches (R1LL and R2LL) only.

a. Reconfigure the lower leaf switch ports (R1LL and R2LL).

Run the following command sequence on both of the lower leaf switches (R1LL and R2LL).

You must use the correct switch configuration file, which you earlier copied to the switch (in step 1.g). In this step, the configuration file name must end with `step5.cfg`.

i. On R1LL:

```

R1LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1LL(config)# run-script bootflash:///roce_multi_online_step5.cfg
| grep 'none'
R1LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1LL(config)# <Ctrl-Z>
R1LL#
  
```

ii. Repeat the command sequence on R2LL:

```

R2LL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2LL(config)# run-script bootflash:///roce_multi_online_step5.cfg
| grep 'none'
R2LL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2LL(config)# <Ctrl-Z>
R2LL#
  
```

 **Note:**

This step can take approximately 5 to 8 minutes on each switch.

- b. On both racks, cable the lower leaf switch to both of the spine switches (R1LL and R2LL to R1SS and R2SS).

Connect the cables from the spine switches that you prepared earlier (in step 2.d).

Cable the switches as described in [Two-Rack Cabling for X8M Racks](#):

- i. On R1LL, cable ports 04, 05, 06, 07, 30, 31, 32, and 33 to R1SS and R2SS.
 - ii. On R2LL, cable ports 04, 05, 06, 07, 30, 31, 32, and 33 to R1SS and R2SS.
- c. On the lower leaf switches, verify that all of the cabled RoCE Network Fabric ports are connected (R1LL and R2LL).

Run the `show interface status` command on each lower leaf switch:

- i. R1LL# `show interface status`

- ii. R2LL# `show interface status`

Examine the output to ensure that all of the cabled ports are connected.

For example:

R1LL# `show interface status`

```

-----
-----
Port          Name          Status  Vlan    Duplex
Speed  Type
-----
-----
mgmt0        --          connected routed    full
1000        --
-----
-----
Port          Name          Status  Vlan    Duplex
Speed  Type
-----
-----
...
Eth1/4        RouterPort1   connected routed    full
100G  QSFP-100G-CR4
Eth1/5        RouterPort2   connected routed    full
100G  QSFP-100G-CR4
Eth1/6        RouterPort3   connected routed    full
100G  QSFP-100G-CR4
Eth1/7        RouterPort4   connected routed    full
100G  QSFP-100G-CR4
Eth1/8        celadm14      connected 3888    full
100G  QSFP-100G-CR4
Eth1/9        celadm13      connected 3888    full
100G  QSFP-100G-CR4
Eth1/10       celadm12      connected 3888    full

```

```

100G    QSFP-100G-CR4
Eth1/11    celadm11    connected 3888    full    100G
QSFP-100G-CR4
Eth1/12    celadm10    connected 3888    full    100G
QSFP-100G-CR4
Eth1/13    celadm09    connected 3888    full    100G
QSFP-100G-CR4
Eth1/14    celadm08    connected 3888    full    100G
QSFP-100G-CR4
Eth1/15    adm08    connected 3888    full    100G
QSFP-100G-CR4
Eth1/16    adm07    connected 3888    full    100G
QSFP-100G-CR4
Eth1/17    adm06    connected 3888    full    100G
QSFP-100G-CR4
Eth1/18    adm05    connected 3888    full    100G
QSFP-100G-CR4
...

```

 **Note:**

Before proceeding, ensure that you have completed all of the actions to this point in step 5 on both lower leaf switches (R1LL and R2LL). If not, then ensure that you go back and perform the missing actions.

- d. Verify the configuration of the lower leaf switches.

You can use the instance of `patchmgr` that you previously used to update the switch firmware (in step 1.h).

Use a switch list file (`l1.lst`) to check both lower leaf switches using one `patchmgr` command:

```

# cat l1.lst
R1LL_IP:mleaf.101
R2LL_IP:mleaf.103

```

On a system with Secure Fabric enabled, use the `msfleaf` tag in the switch list file:

```

# cat l1.lst
R1LL_IP:msfleaf.101
R2LL_IP:msfleaf.103

```

The following shows the recommended command and an example of the expected results:

```

# ./patchmgr --roceswitches l1.lst --verify-config -log_dir /tmp/log
2020-08-10 13:45:09 -0700          :Working: Initiating config
verification... Expect up to 6 minutes for each switch
Mon Aug 10 13:45:13 PDT 2020 1 of 4 :Verifying config on switch ...
...
Mon Aug 10 13:45:32 PDT 2020:          [INFO          ] Config matches

```



```

template: ...
Mon Aug 10 13:45:32 PDT 2020:      [SUCCESS ] Config validation
successful!
2020-08-10 13:45:32 -0700
  Config check on RoCE switch(es)
2020-08-10 13:45:32 -0700
  Completed run of command: ./patchmgr --roceswitches ll.lst --
verify-config -log_dir /tmp/log
2020-08-10 13:45:32 -0700      :INFO      : config attempted on
nodes in file ll.lst: [R1LL_IP R2LL_IP]
2020-08-10 13:45:32 -0700      :INFO      : For details, check
the following files in /tmp/log:
2020-08-10 13:45:32 -0700      :INFO      : -
updateRoceSwitch.log
2020-08-10 13:45:32 -0700      :INFO      : -
updateRoceSwitch.trc
2020-08-10 13:45:32 -0700      :INFO      : - patchmgr.stdout
2020-08-10 13:45:32 -0700      :INFO      : - patchmgr.stderr
2020-08-10 13:45:32 -0700      :INFO      : - patchmgr.log
2020-08-10 13:45:32 -0700      :INFO      : - patchmgr.trc
2020-08-10 13:45:32 -0700      :INFO      : Exit status:0
2020-08-10 13:45:32 -0700      :INFO      : Exiting.
  
```

In the command output, verify that the switch configuration is good for both lower leaf switches.

- e. Verify that nve is up on the lower leaf switches (R1LL and R2LL).

Run the following command on each lower leaf switch and examine the output:

- i. R1LL# **show nve peers**
- ii. R2LL# **show nve peers**

At this point, you should see one nve peer with State=Up.

For example:

```

R1LL# show nve peers
Interface      Peer-IP State LearnType  Uptime      Router-
Mac
-----
-----
nve1           100.64.1.103  Up      CP 00:04:29
n/a
  
```

- f. Verify that BGP is up on the lower leaf switches (R1LL and R2LL).

Run the following command on each lower leaf switch and examine the output:

- i. R1LL# **show logging log | grep BGP**
- ii. R2LL# **show logging log | grep BGP**

Look for two entries with `Up` in the rightmost column that are associated with different IP addresses.

For example:

```
R1LL# show logging log | grep BGP
2020 Aug 10 13:47:13 R1LL %BGP-5-ADJCHANGE: bgp- [29342] (default)
neighbor 100.64.0.201 Up
2020 Aug 10 13:47:24 R1LL %BGP-5-ADJCHANGE: bgp- [29342] (default)
neighbor 100.64.0.202 Up
```

6. Finalize the configuration of the upper leaf switches (R1UL and R2UL).

Perform this step on the upper leaf switches (R1UL and R2UL) only.

a. Start the inter-switch ports on the upper leaf switches (R1UL and R2UL).

i. On R1UL:

```
R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/4-7, ethernet 1/30-33
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
```

ii. Repeat the command sequence on R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/4-7, ethernet 1/30-33
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

b. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.

c. Verify the status of the inter-switch ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

i. R1UL# `show interface status`

ii. R2UL# `show interface status`

Examine the output to ensure that the inter-switch ports are connected.

For example:

R1UL# **show interface status**

```

-----
-----
Port          Name          Status      Vlan      Duplex
Speed   Type
-----
-----
mgmt0          --          connected  routed    full
1000          --
-----
-----
Port          Name          Status      Vlan      Duplex
Speed   Type
-----
-----
Eth1/1          --          xcvrAbsen  1          auto
auto          --
Eth1/2          --          xcvrAbsen  1          auto
auto          --
Eth1/3          --          xcvrAbsen  1          auto
auto          --
Eth1/4          ISL1          connected  routed    full
100G   QSFP-100G-CR4
Eth1/5          ISL2          connected  routed    full
100G   QSFP-100G-CR4
Eth1/6          ISL3          connected  routed    full
100G   QSFP-100G-CR4
Eth1/7          ISL4          connected  routed    full
100G   QSFP-100G-CR4
Eth1/8          celadm14      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/9          celadm13      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/10         celadm12      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/11         celadm11      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/12         celadm10      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/13         celadm09      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/14         celadm08      disabled   3888      full
100G   QSFP-100G-CR4
Eth1/15         adm08         disabled   3888      full
100G   QSFP-100G-CR4
Eth1/16         adm07         disabled   3888      full
100G   QSFP-100G-CR4
Eth1/17         adm06         disabled   3888      full
100G   QSFP-100G-CR4
    
```

```

Eth1/18      adm05      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/19      adm04      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/20      adm03      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/21      adm02      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/22      adm01      disabled 3888      full 100G
QSFP-100G-CR4
Eth1/23      celadm07   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/24      celadm06   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/25      celadm05   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/26      celadm04   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/27      celadm03   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/28      celadm02   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/29      celadm01   disabled 3888      full 100G
QSFP-100G-CR4
Eth1/30      ISL5       connected routed    full 100G
QSFP-100G-CR4
Eth1/31      ISL6       connected routed    full 100G
QSFP-100G-CR4
Eth1/32      ISL7       connected routed    full 100G
QSFP-100G-CR4
Eth1/33      ISL8       connected routed    full 100G
QSFP-100G-CR4
Eth1/34      --         xcvrAbsen 1        auto auto  --
Eth1/35      --         xcvrAbsen 1        auto auto  --
Eth1/36      --         xcvrAbsen 1        auto auto  --
Po100       --         connected trunk    full 100G  --
Lo0         Routing loopback i connected routed    auto auto  --
Lo1         VTEP loopback inte connected routed    auto auto  --
Vlan1      --         down routed        auto auto  --
nve1       --         connected --        auto auto  --
  
```

d. Start the storage server ports on the upper leaf switches (R1UL and R2UL).

i. On R1UL:

```

R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
  
```

```
R1UL(config)# <Ctrl-Z>
R1UL#
```

- ii. Repeat the command sequence on R2UL:

```
R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/8-14, ethernet 1/23-29
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
```

- e. Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.
- f. Verify the status of the storage server ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

- i. R1UL# **show interface status**

- ii. R2UL# **show interface status**

Examine the output to ensure that the storage server ports are connected.

For example:

```
R1UL# show interface status
```

Port	Name	Status	Vlan	Duplex
mgmt0	--	connected	routed	full
1000	--			

Port	Name	Status	Vlan	Duplex
Eth1/1	--	xcvrAbsen	1	auto
auto	--			
Eth1/2	--	xcvrAbsen	1	auto
auto	--			
Eth1/3	--	xcvrAbsen	1	auto
auto	--			

Eth1/4	ISL1	connected routed	full	100G
QSF-100G-CR4				
Eth1/5	ISL2	connected routed	full	100G
QSF-100G-CR4				
Eth1/6	ISL3	connected routed	full	100G
QSF-100G-CR4				
Eth1/7	ISL4	connected routed	full	100G
QSF-100G-CR4				
Eth1/8	celadm14	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/9	celadm13	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/10	celadm12	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/11	celadm11	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/12	celadm10	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/13	celadm09	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/14	celadm08	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/15	adm08	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/16	adm07	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/17	adm06	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/18	adm05	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/19	adm04	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/20	adm03	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/21	adm02	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/22	adm01	disabled 3888	full	100G
QSF-100G-CR4				
Eth1/23	celadm07	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/24	celadm06	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/25	celadm05	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/26	celadm04	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/27	celadm03	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/28	celadm02	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/29	celadm01	<i>connected</i> 3888	full	100G
QSF-100G-CR4				
Eth1/30	ISL5	connected routed	full	100G
QSF-100G-CR4				
Eth1/31	ISL6	connected routed	full	100G

```

QSF1P-100G-CR4
Eth1/32      ISL7          connected routed    full
100G        QSF1P-100G-CR4
Eth1/33      ISL8          connected routed    full
100G        QSF1P-100G-CR4
Eth1/34      --            xcvrAbsen 1       auto
auto        --
Eth1/35      --            xcvrAbsen 1       auto
auto        --
Eth1/36      --            xcvrAbsen 1       auto
auto        --
Po100        --            connected trunk   full
100G        --
Lo0          Routing loopback i connected routed    auto
auto        --
Lo1          VTEP loopback inte connected routed    auto
auto        --
Vlan1        --            down      routed    auto
auto        --
nve1         --            connected --      auto
auto        --
  
```

- g.** Start the database server ports on the upper leaf switches (R1UL and R2UL).

- i.** On R1UL:

```

R1UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R1UL(config)# interface ethernet 1/15-22
R1UL(config-if-range)# no shut
R1UL(config-if-range)# exit
R1UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R1UL(config)# <Ctrl-Z>
R1UL#
  
```

- ii.** Repeat the command sequence on R2UL:

```

R2UL# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R2UL(config)# interface ethernet 1/15-22
R2UL(config-if-range)# no shut
R2UL(config-if-range)# exit
R2UL(config)# copy running-config startup-config
[#####] 100%
Copy complete, now saving to disk (please wait)...
Copy complete
R2UL(config)# <Ctrl-Z>
R2UL#
  
```

- h.** Wait for 5 minutes to ensure that the ports you just started are fully operational before continuing.

- i. Verify the status of the database server ports on the upper leaf switches (R1UL and R2UL).

Run the `show interface status` command on each upper leaf switch:

- i. R1UL# `show interface status`

- ii. R2UL# `show interface status`

Examine the output to ensure that the database server ports are connected.

For example:

R1UL# `show interface status`

```

-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed
Type
-----
-----
Eth1/1        --            xcvrAbsen  1         auto    auto    --
Eth1/2        --            xcvrAbsen  1         auto    auto    --
Eth1/3        --            xcvrAbsen  1         auto    auto    --
Eth1/4        ISL1          connected   routed    full    100G
QSFP-100G-CR4
Eth1/5        ISL2          connected   routed    full    100G
QSFP-100G-CR4
Eth1/6        ISL3          connected   routed    full    100G
QSFP-100G-CR4
Eth1/7        ISL4          connected   routed    full    100G
QSFP-100G-CR4
Eth1/8        celadm14     connected   3888     full    100G
QSFP-100G-CR4
Eth1/9        celadm13     connected   3888     full    100G
QSFP-100G-CR4
Eth1/10       celadm12     connected   3888     full    100G
QSFP-100G-CR4
Eth1/11       celadm11     connected   3888     full    100G
QSFP-100G-CR4
Eth1/12       celadm10     connected   3888     full    100G
QSFP-100G-CR4
Eth1/13       celadm09     connected   3888     full    100G
QSFP-100G-CR4
Eth1/14       celadm08     connected   3888     full    100G
QSFP-100G-CR4
Eth1/15       adm08        connected   3888     full    100G

```



```

QSF1P-100G-CR4
Eth1/16      adm07      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/17      adm06      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/18      adm05      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/19      adm04      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/20      adm03      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/21      adm02      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/22      adm01      connected 3888      full
100G      QSF1P-100G-CR4
Eth1/23      celadm07   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/24      celadm06   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/25      celadm05   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/26      celadm04   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/27      celadm03   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/28      celadm02   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/29      celadm01   connected 3888      full
100G      QSF1P-100G-CR4
Eth1/30      ISL5       connected routed    full
100G      QSF1P-100G-CR4
Eth1/31      ISL6       connected routed    full
100G      QSF1P-100G-CR4
Eth1/32      ISL7       connected routed    full
100G      QSF1P-100G-CR4
Eth1/33      ISL8       connected routed    full
100G      QSF1P-100G-CR4
Eth1/34      --         xcvrAbsen 1        auto
auto      --
Eth1/35      --         xcvrAbsen 1        auto
auto      --
Eth1/36      --         xcvrAbsen 1        auto
auto      --
Po100      --         connected trunk    full
100G      --
Lo0        Routing loopback i connected routed    auto
auto      --
Lo1        VTEP loopback inte connected routed    auto
auto      --
Vlan1      --         down      routed    auto
auto      --
nve1      --         connected --        auto
auto      --
  
```

- j. Verify that nve is up on the leaf switches (R1LL, R1UL, R2LL, and R2UL).

Run the following command on each leaf switch and examine the output:

i. R1LL# **show nve peers**

ii. R1UL# **show nve peers**

iii. R2LL# **show nve peers**

iv. R2UL# **show nve peers**

In the output, you should see three nve peers with `State=Up`.

For example:

```
R1UL# show nve peers
Interface          Peer-IP State LearnType  Uptime      Router-Mac
-----
nve1               100.64.1.101 Up        CP 00:04:29  n/a
nve1               100.64.1.103 Up        CP 00:07:48  n/a
nve1               100.64.1.104 Up        CP 00:04:10  n/a
```

- k. Verify that BGP is up on the upper leaf switches (R1UL and R2UL).

Run the following command on each upper leaf switch and examine the output:

i. R1UL# **show logging log | grep BGP**

ii. R2UL# **show logging log | grep BGP**

In the output, look for two entries with `Up` in the rightmost column that are associated with different IP addresses.

For example:

```
R1UL# show logging log | grep BGP
2020 Aug 10 13:57:13 R1UL %BGP-5-ADJCHANGE: bgp- [32782] (default)
neighbor 100.64.0.201 Up
2020 Aug 10 13:57:24 R1UL %BGP-5-ADJCHANGE: bgp- [32782] (default)
neighbor 100.64.0.202 Up
```

7. For each rack (R1 and R2), confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

The `verify_roce_cables.py` script uses two input files; one for database servers and storage servers (`nodes.rackN`), and another for switches (`switches.rackN`). In each file, every server or switch must be listed on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

See My Oracle Support document 2587717.1 for download and detailed usage instructions.

Run the `verify_roce_cables.py` script against both of the racks:

- a. # **cd /opt/oracle.SupportTools/ibdiagtools**
./verify_roce_cables.py -n nodes.rack1 -s switches.rack1

```
b. # cd /opt/oracle.SupportTools/ibdiagtools
   # ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
```

Check the output of the `verify_roce_cables.py` script against the tables in [Two-Rack Cabling for X8M Racks](#). Also, check that output in the `CABLE OK?` columns contains the `OK` status.

The following examples show extracts of the expected command results:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER) LOWER LEAF (rack1sw-rocea0) : CABLE OK? UPPER LEAF
(rack1sw-roceb0) : CABLE OK?
-----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/17 : OK rack1sw-
roces0 Ethernet1/9 : OK
Eth1/5 (ISL peer switch) : rack1sw-roces0 Ethernet1/13 : OK rack1sw-
roces0 Ethernet1/5 : OK
Eth1/6 (ISL peer switch) : rack1sw-roces0 Ethernet1/19 : OK rack1sw-
roces0 Ethernet1/11 : OK
Eth1/7 (ISL peer switch) : rack1sw-roces0 Ethernet1/15 : OK rack1sw-
roces0 Ethernet1/7 : OK
Eth1/12 (celadm10) : rack1celadm10 port-1 : OK
rack1celadm10 port-2 : OK
Eth1/13 (celadm09) : rack1celadm09 port-1 : OK
rack1celadm09 port-2 : OK
Eth1/14 (celadm08) : rack1celadm08 port-1 : OK
rack1celadm08 port-2 : OK
...
Eth1/15 (adm08) : rack1dbadm08 port-1 : OK rack1dbadm08
port-2 : OK
Eth1/16 (adm07) : rack1dbadm07 port-1 : OK rack1dbadm07
port-2 : OK
Eth1/17 (adm06) : rack1dbadm06 port-1 : OK rack1dbadm06
port-2 : OK
...
Eth1/30 (ISL peer switch) : rack2sw-roces0 Ethernet1/17 : OK rack2sw-
roces0 Ethernet1/9 : OK
Eth1/31 (ISL peer switch) : rack2sw-roces0 Ethernet1/13 : OK rack2sw-
roces0 Ethernet1/5 : OK
Eth1/32 (ISL peer switch) : rack2sw-roces0 Ethernet1/19 : OK rack2sw-
roces0 Ethernet1/11 : OK
Eth1/33 (ISL peer switch) : rack2sw-roces0 Ethernet1/15 : OK rack2sw-
roces0 Ethernet1/7 : OK

# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
SWITCH PORT (EXPECTED PEER) LOWER LEAF (rack2sw-rocea0) : CABLE OK? UPPER LEAF
(rack2sw-roceb0) : CABLE OK?
-----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/18 : OK rack1sw-
```

```
roces0 Ethernet1/10 : OK
...
```

8. Verify the RoCE Network Fabric operation across both interconnected racks by using the `infinicheck` command.

Use the following recommended command sequence to verify the RoCE Network Fabric operation across both racks.

In each command, `hosts.all` contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses from both racks (2 RoCE Network Fabric IP addresses for each database server), and `cells.all` contains a list of RoCE Network Fabric IP addresses for the storage servers from both racks (2 RoCE Network Fabric IP addresses for each storage server).

- a.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.all -c cells.all -z
```
- b.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.all -c cells.all -s
```
- c.

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.all -c cells.all -b
```

See step 1.k for most information about each `infinicheck` command.

The following example shows the expected command results for the final command in the sequence:

```
# cd /opt/oracle.SupportTools/ibdiagtools
# ./infinicheck -g hosts.all -c cells.all -b

INFINICHECK
      [Network Connectivity, Configuration and Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
      ##### ROCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS]... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS]... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS]... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs and CELLS
[SUCCESS]... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value : 2300
[SUCCESS]... RoCE interface MTU settings look good
      Verifying switch advertised DSCP on all DBs and CELLS ports ( )
[SUCCESS]... Advertised DSCP settings from RoCE switch looks good
      ##### CONNECTIVITY TESTS #####
      [COMPUTE NODES -> STORAGE CELLS]
      (60 seconds approx.)
      (Will walk through QoS values: 0-6) [SUCCESS].....Results OK
```

```
[SUCCESS]..... All can talk to all storage cells  
[COMPUTE NODES -> COMPUTE NODES]  
...
```

At this point, both racks share the RoCE Network Fabric, and the combined system is ready for further configuration. See [Configuring the New Hardware](#).

Related Topics

- [Cabling Two Oracle Exadata Database Machine RoCE Network Fabric Racks With No Downtime \(My Oracle Support Doc ID 2704997.1\)](#)
- [Verify RoCE Cabling on Oracle Exadata Database Machine X8M-2 and X8M-8 Servers \(My Oracle Support Doc ID 2587717.1\)](#)

2.3.2.2 Cabling Two RoCE Network Fabric Racks Together with Down Time

If your operational requirements can tolerate some scheduled down time, then you can choose from the following procedures to extend your existing RoCE Network Fabric rack by adding another rack.

- [Cabling Two RoCE Network Fabric Racks Together with Down Time using Oracle Exadata System Software Release 20.1.0 or Later](#)
Use this simpler procedure to cable together two racks with RoCE Network Fabric where some down-time can be tolerated and you are using Oracle Exadata System Software release 20.1.0, or later.
- [Cabling Two RoCE Network Fabric Racks Together with Down Time using Oracle Exadata System Software Release 19.3 or Earlier](#)
Use this procedure to cable together two racks with RoCE Network Fabric where some down-time can be tolerated and you are using Oracle Exadata System Software Release 19.3, or earlier.

2.3.2.2.1 Cabling Two RoCE Network Fabric Racks Together with Down Time using Oracle Exadata System Software Release 20.1.0 or Later

Use this simpler procedure to cable together two racks with RoCE Network Fabric where some down-time can be tolerated and you are using Oracle Exadata System Software release 20.1.0, or later.

In this procedure, the existing rack is R1, and the new rack is R2.

Use the applicable cabling tables depending on your system configuration:

- [Two-Rack Cabling for X9M Racks](#)
- [Two-Rack Cabling for X8M Racks](#)
- [Two-Rack Cabling for a System Combining X8M and X9M Racks](#)

1. Ensure the new rack is near the existing rack.

The RDMA Network Fabric cables must be able to reach the servers in each rack.

2. Ensure you have a backup of the current switch configuration for each switch in the existing and new rack.

See [Backing Up Settings on the RoCE Network Fabric Switch](#) in *Oracle Exadata Database Machine Maintenance Guide*.

3. Shut down all servers on both the new rack (R2) and the existing rack (R1).

The switches should remain available.

4. Update the firmware to the latest available release on all of the RoCE Network Fabric switches.

For this step, treat all of the switches as if they belong to a single rack system.

See Updating RoCE Network Fabric Switch Firmware in *Oracle Exadata Database Machine Maintenance Guide*.

5. Apply the multi-rack golden configuration settings on the RoCE Network Fabric switches.

Use the procedure described in Applying Golden Configuration Settings on RoCE Network Fabric Switches, in *Oracle Exadata Database Machine Maintenance Guide*.

6. Perform the physical cabling of the switches.
 - a. In Rack 2, remove the existing inter-switch connections between the two leaf switches, R2UL and R2LL.
 - b. In Rack 2, cable each leaf switch to the spine switches using the applicable cabling tables.
 - c. In Rack 1, remove the existing inter-switch connections between the two leaf switches, R1UL and R1LL.
 - d. In Rack 1, cable each leaf switch to the spine switches using the applicable cabling tables.
7. Confirm each switch is available and connected.

For each of the 6 switches, confirm the output from the `show interface status` command shows `connected` and `100G` for each connected inter-switch port. Use the appropriate cabling tables to identify the ports that should be connected.

In the following examples, the leaf switches are ports Eth1/4 to Eth1/7, and Eth1/30 to Eth1/33. The spine switches are ports Eth1/5 to Eth1/20.

When run from a spine switch, the output should be similar to the following:

```
rack1sw-roces0# show interface status
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
-----
...
Eth1/5       RouterPort5   connected  routed    full    100G
QSFP-100G-CR4
Eth1/6       RouterPort6   connected  routed    full    100G
QSFP-100G-SR4
Eth1/7       RouterPort7   connected  routed    full    100G
QSFP-100G-CR4
Eth1/8       RouterPort8   connected  routed    full    100G
QSFP-100G-SR4
Eth1/9       RouterPort9   connected  routed    full    100G
```

```

QSF-P-100G-CR4
Eth1/10      RouterPort10      connected routed    full
100G QSF-P-100G-SR4
Eth1/11      RouterPort11      connected routed    full
100G QSF-P-100G-CR4
Eth1/12      RouterPort12      connected routed    full
100G QSF-P-100G-SR4
Eth1/13      RouterPort13      connected routed    full
100G QSF-P-100G-CR4
Eth1/14      RouterPort14      connected routed    full
100G QSF-P-100G-SR4
Eth1/15      RouterPort15      connected routed    full
100G QSF-P-100G-CR4
Eth1/16      RouterPort16      connected routed    full
100G QSF-P-100G-SR4
Eth1/17      RouterPort17      connected routed    full
100G QSF-P-100G-CR4
Eth1/18      RouterPort18      connected routed    full
100G QSF-P-100G-SR4
Eth1/19      RouterPort19      connected routed    full
100G QSF-P-100G-CR4
Eth1/20      RouterPort20      connected routed    full
100G QSF-P-100G-SR4
Eth1/21      RouterPort21      xcvrAbsen    routed    full
100G      --
...
  
```

When run from a leaf switch, the output should be similar to the following:

```

rack1sw-rocea0# show interface status
-----
-----
Port      Name      Status  Vlan  Duplex
Speed    Type
-----
-----
mgmt0     --        connected routed  full
1000     --
-----
-----
Port      Name      Status  Vlan  Duplex
Speed    Type
-----
-----
...
Eth1/4    RouterPort1 connected routed  full
100G QSF-P-100G-CR4
Eth1/5    RouterPort2 connected routed  full
100G QSF-P-100G-CR4
Eth1/6    RouterPort3 connected routed  full
100G QSF-P-100G-CR4
Eth1/7    RouterPort4 connected routed  full
100G QSF-P-100G-CR4
Eth1/8    celadm14    connected 3888    full
  
```

```

100G    QSFP-100G-CR4
...
Eth1/29    celadm01    connected 3888    full    100G
QSFP-100G-CR4
Eth1/30    RouterPort5    connected routed    full    100G
QSFP-100G-SR4
Eth1/31    RouterPort6    connected routed    full    100G
QSFP-100G-SR4
Eth1/32    RouterPort7    connected routed    full    100G
QSFP-100G-SR4
Eth1/33    RouterPort8    connected routed    full    100G
QSFP-100G-SR4
...

```

8. Check the neighbor discovery for every switch in racks R1 and R2.

Log in to each switch and use the `show lldp neighbors` command. Make sure that all switches are visible and check the switch ports assignment against the applicable cabling tables.

A spine switch should see the two leaf switches in each rack, but not the other spine switch. The output for a spine switch should be similar to the following:

 **Note:**

The interfaces output in the `Port ID` column are different for each switch based on the applicable cabling tables.

```

rack1sw-roces0# show lldp neighbors
...
Device ID           Local Intf           Hold-time  Capability  Port ID
rack1-adm0          mgmt0                120       BR          Ethernet1/47
rack1sw-roceb0     Eth1/5             120       BR          Ethernet1/5
rack2sw-roceb0     Eth1/6             120       BR          Ethernet1/5
rack1sw-roceb0     Eth1/7             120       BR          Ethernet1/7
rack2sw-roceb0     Eth1/8             120       BR          Ethernet1/7
rack1sw-roceb0     Eth1/9             120       BR          Ethernet1/4
rack2sw-roceb0     Eth1/10            120       BR          Ethernet1/4
rack1sw-roceb0     Eth1/11            120       BR          Ethernet1/6
rack2sw-roceb0     Eth1/12            120       BR          Ethernet1/6
rack1sw-rocea0     Eth1/13            120       BR          Ethernet1/5
rack2sw-rocea0     Eth1/14            120       BR          Ethernet1/5
rack1sw-rocea0     Eth1/15            120       BR          Ethernet1/7
rack2sw-rocea0     Eth1/16            120       BR          Ethernet1/7
rack1sw-rocea0     Eth1/17            120       BR          Ethernet1/4
rack2sw-rocea0     Eth1/18            120       BR          Ethernet1/4
rack1sw-rocea0     Eth1/19            120       BR          Ethernet1/6
rack2sw-rocea0     Eth1/20            120       BR          Ethernet1/6
Total entries displayed: 17

```

Each leaf switch should see the two spine switches, but not the other leaf switches. The output for a leaf switch should be similar to the following:

 **Note:**

The interfaces output in the `Port ID` column are different for each switch based on the applicable cabling tables.

```
rack1sw-rocea0# show lldp neighbors
...
Device ID           Local Intf      Hold-time  Capability  Port ID
switch              mgmt0          120        BR          Ethernet1/46
rack1sw-roces0     Eth1/4         120        BR          Ethernet1/17
rack1sw-roces0     Eth1/5         120        BR          Ethernet1/13
rack1sw-roces0     Eth1/6         120        BR          Ethernet1/19
rack1sw-roces0     Eth1/7         120        BR          Ethernet1/15
rack2sw-roces0     Eth1/30        120        BR          Ethernet1/17
rack2sw-roces0     Eth1/31        120        BR          Ethernet1/13
rack2sw-roces0     Eth1/32        120        BR          Ethernet1/19
rack2sw-roces0     Eth1/33        120        BR          Ethernet1/15
rocetoi-ext-sw     Eth1/36        120        BR          Ethernet1/49
Total entries displayed: 10
```

9. Power on all servers in racks R1 and R2.
10. For each rack, confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

Refer to My Oracle Support Doc ID 2587717.1 for download and usage instructions.

Check the output of the `verify_roce_cables.py` script against the applicable cabling tables. Also, check that output in the `CABLE OK?` columns contains the `OK` status.

When running the script, two input files are used, one for nodes and one for switches. Each file should contain the servers or switches on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

The following output is a partial example of the command results:

```
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER) LEAF-1 (rack1sw-rocea0) : CABLE OK? LEAF-2
(rack1sw-roceb0) : CABLE OK?
-----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/17 : OK rack1sw-
roces0 Ethernet1/9 : OK
Eth1/5 (ISL peer switch) : rack1sw-roces0 Ethernet1/13 : OK rack1sw-
roces0 Ethernet1/5 : OK
Eth1/6 (ISL peer switch) : rack1sw-roces0 Ethernet1/19 : OK rack1sw-
roces0 Ethernet1/11: OK
Eth1/7 (ISL peer switch) : rack1sw-roces0 Ethernet1/15 : OK rack1sw-
roces0 Ethernet1/7 : OK
Eth1/12 (celadm10) : rack1celadm10 port-1 : OK
rack1celadm10 port-2 : OK
Eth1/13 (celadm09) : rack1celadm09 port-1 : OK
rack1celadm09 port-2 : OK
```

```

Eth1/14 (celadm08)      : rack1celadm08 port-1      : OK      rack1celadm08
port-2      : OK
...
Eth1/15 (adm08)        : rack1dbadm08 port-1      : OK      rack1dbadm08
port-2      : OK
Eth1/16 (adm07)        : rack1dbadm07 port-1      : OK      rack1dbadm07
port-2      : OK
Eth1/17 (adm06)        : rack1dbadm06 port-1      : OK      rack1dbadm06
port-2      : OK
...
Eth1/30 (ISL peer switch) : rack2sw-roces0 Ethernet1/17 : OK      rack2sw-roces0
Ethernet1/9 : OK
Eth1/31 (ISL peer switch) : rack2sw-roces0 Ethernet1/13 : OK      rack2sw-roces0
Ethernet1/5 : OK
Eth1/32 (ISL peer switch) : rack2sw-roces0 Ethernet1/19 : OK      rack2sw-roces0
Ethernet1/11: OK
Eth1/33 (ISL peer switch) : rack2sw-roces0 Ethernet1/15 : OK      rack2sw-roces0
Ethernet1/7 : OK

# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
SWITCH PORT (EXPECTED PEER) LEAF-1 (rack2sw-rocea0) : CABLE OK? LEAF-2 (rack2sw-
roceb0) : CABLE OK?
-----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/18 : OK      rack1sw-roces0
Ethernet1/10: OK
...

```

11. Verify the RoCE Network Fabric operation by using the `infinicheck` command.

Use the following recommended command sequence. In each command, `hosts.lst` is the name of an input file that contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses, and `cells.lst` is the name of an input file that contains a list of RoCE Network Fabric IP addresses for the storage servers.

- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -z
```

- Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -s
```

- Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -b
```

```

INFINICHECK
  [Network Connectivity, Configuration and
  Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all
CELLs.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLs
[SUCCESS]... RoCE devices on all DBs and CELLs look good
      Checking for RoCE Policy Routing settings on all DBs and
CELLs
[SUCCESS]... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLs
[SUCCESS]... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs
and CELLs
[SUCCESS]... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value :
2300
[SUCCESS]... RoCE interface MTU settings look good
      Verifying switch advertised DSCP on all DBs and CELLs ports
( )
[SUCCESS]... Advertised DSCP settings from RoCE switch looks
good
      ##### CONNECTIVITY TESTS #####
      [COMPUTE NODES -> STORAGE CELLS]
      (60 seconds approx.)
      (Will walk through QoS values: 0-6)
[SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
      [COMPUTE NODES -> COMPUTE NODES]
...

```

12. After cabling the racks together, proceed to [Configuring the New Hardware](#) to finish the configuration of the new rack.

Related Topics

- [Verify RoCE Cabling on Oracle Exadata Database Machine X8M-2 and X8M-8 Servers \(My Oracle Support Doc ID 2587717.1\)](#)

2.3.2.2.2 Cabling Two RoCE Network Fabric Racks Together with Down Time using Oracle Exadata System Software Release 19.3 or Earlier

Use this procedure to cable together two racks with RoCE Network Fabric where some down-time can be tolerated and you are using Oracle Exadata System Software Release 19.3, or earlier.

In this procedure, the existing rack is R1, and the new rack is R2.

1. Ensure the new rack is near the existing rack.

The RDMA Network Fabric cables must be able to reach the servers in each rack.

2. Ensure you have a backup of the current switch configuration for each switch in the existing and new rack.

See Backing Up Settings on the RoCE Network Fabric Switch in *Oracle Exadata Database Machine Maintenance Guide*.

3. Shut down all servers on both the new rack (R2) and the existing rack (R1).

The switches should remain available.

4. Update the firmware to the latest available release on all of the RoCE Network Fabric switches.

For this step, treat all of the switches as if they belong to a single rack system.

See Updating RoCE Network Fabric Switch Firmware in *Oracle Exadata Database Machine Maintenance Guide*.

5. Apply the multi-rack spine switch configuration to the two spine switches.

- a. Log in to the server that has downloaded the latest RDMA Network Fabric patch ZIP file.

To find the available RDMA Network Fabric patches, search for '**RDMA network switch**' in My Oracle Support document 888828.1. Download and use the latest patch for your Oracle Exadata System Software release.

- b. Unzip the RDMA Network Fabric patch ZIP file and change directories to the location of the `patchmgr` utility.

- c. Make a copy of the golden configuration file for each switch.

Run these commands from patch directory:

```
# cp roce_switch_templates/roce_spine_switch_multi.cfg
roce_spine_switch_multi_R1SS.cfg
# cp roce_switch_templates/roce_spine_switch_multi.cfg
roce_spine_switch_multi_R2SS.cfg
```

- d. Edit each copy of the spine switch configuration file.

Using a text editor, replace each occurrence of `%SPINE_LOOPBACK_IP0%` with the correct IP address for the switch, as indicated in the table below.

Switch	SPINE_LOOPBACK_IP0
Rack 1 spine switch (R1SS)	100.64.0.201
Rack 2 spine switch (R2SS)	100.64.0.202

The following example shows using IP address 100.64.0.201 as the `SPINE_LOOPBACK_IP0` for the spine switch in rack 1 (R1SS).

```
! Define loopback interface for underlay OSPF routing
interface loopback0
  description Routing loopback interface
  !ip address 100.64.0.201/32
  ip address 100.64.0.201/32
  ip router ospf UNDERLAY area 0.0.0.0

! Configure OSPF as the underlay network
router ospf UNDERLAY
```

```

router-id 100.64.0.201

! change ECMP hash rotate value from default 32 to 40 for better
! router port utilization for upto parallel flows via the 8
! available router ports
ip load-sharing address source-destination port source-
destination rotate 40

! Create BGP route reflector to exchange routes across VTEPs
! Use CIDR block of IPs for neighbor range
! - log-neighbor-changes: Enables the generation of logging
messages
! generated when the status of a BGP neighbor changes.
! - address-family ipv4 unicast: Enters address family
configuration
! mode and Specifies IP Version 4 unicast address prefixes.
! address
router bgp 65502
  router-id 100.64.0.201
  log-neighbor-changes

```

- e. Verify the three replacements in the spine switch configuration file.

For example, for rack 1 (R1SS) check for IP address 100.64.0.201 in the spine switch configuration file:

```

$ grep 100.64 roce_spine_switch_multi_R1SS.cfg |grep -v
'neighbor' |grep -v '! '
ip address 100.64.0.201/32
router-id 100.64.0.201
router-id 100.64.0.201

```

- f. Apply the updated multi-rack configuration file to its corresponding spine switch.
- i. Log in to each spine switch, and remove the existing configuration file using the following command:

```
delete bootflash:roce_spine_switch_multi.cfg
```

For example:

```

rack1sw-roces0# delete bootflash:roce_spine_switch_multi.cfg
Do you want to delete "/roce_spine_switch_multi.cfg" ?
(yes/no/abort) [y] y
rack1sw-roces0#

```

- ii. Log in to the server that contains the modified configuration files, and copy each file to its corresponding spine switch.

```

# scp roce_spine_switch_multi_R1SS.cfg
admin@R1SS_IP_Address:/
# scp roce_spine_switch_multi_R2SS.cfg
admin@R2SS_IP_Address:/

```

- iii. Verify the modified file was copied successfully to the spine switch. For example, for the first rack, log in to the spine switch again and use the following command:

```
rack1sw-roces0# dir bootflash:roce_spine_switch_multi_R1SS.cfg
      27360 Nov 20 12:12:50 2019 roce_spine_switch_multi_R1SS.cfg
Usage for bootflash://sup-local
1829572608 bytes used
114893496320 bytes free
116723068928 bytes total
```

- iv. Log in to each switch again, and copy the modified configuration into flash. On the spine switch for rack 1, you would use the following commands:

```
rack1sw-roces0# run-script
bootflash:roce_spine_switch_multi_R1SS.cfg | grep 'none'
rack1sw-roces0# copy running-config startup-config
```

On the spine switch for rack 2, you would use the following commands:

```
rack2sw-roces0# run-script
bootflash:roce_spine_switch_multi_R2SS.cfg | grep 'none'
rack2sw-roces0# copy running-config startup-config
```

 **Note:**

The `run-script` command for a spine switch can take approximately 2 minutes to complete.

- 6. Apply the multi-rack leaf switch configuration to the four leaf switches.

For each switch, complete the following steps, where *SW#* represents the values *R1LL*, *R1UL*, *R2LL*, or *R2UL*, depending on which switch you are configuring.

- a. Log in to the server that has downloaded the RDMA Network Fabric patch ZIP file (from Step 5.a) for the Oracle Exadata System Software release used by the existing racks.
- b. Change directories to the location of the `patchmgr` utility.
- c. Make a copy of the golden configuration file for each leaf switch.

You can copy either the `roce_leaf_switch_multi.cfg` file or the `roce_qinq_leaf_switch_multi.cfg` file if you want to enable Exadata Secure RDMA Fabric Isolation.

After extracting the patch ZIP file, run the following command four times from the patch directory, substituting for *SW#* the values *R1LL*, *R1UL*, *R2LL*, and *R2UL*.

```
# cp roce_switch_templates/roce_leaf_switch_multi.cfg
roce_leaf_switch_multi_SW#.cfg
```

- d. Edit each copy of the leaf switch configuration file to replace the loopback IP addresses.

Using a text editor, replace each occurrence of `%LEAF_LOOPBACK_IP0%` and `%LEAF_LOOPBACK_IP1%` with the correct IP addresses for the leaf switch, as indicated in the table below.

The scheme used for loopback IP addresses for the leaf switches in a 2-rack system is:

Switch	LEAF_LOOPBACK_IP0	LEAF_LOOPBACK_IP1
Rack 1 Lower Leaf switch (R1LL)	100.64.0.101	100.64.1.101
Rack 1 Upper Leaf switch (R1UL)	100.64.0.102	100.64.1.102
Rack 2 Lower Leaf switch (R2LL)	100.64.0.103	100.64.1.103
Rack 2 Upper Leaf switch (R2UL)	100.64.0.104	100.64.1.104

The following example shows using IP address 100.64.0.101 as the `SPINE_LOOPBACK_IP0` for the lower leaf switch in rack 1 (R1LL).

```
! Define loopback interface for IGP protocol for VTEP
reachability
interface loopback0
  description Routing loopback interface
  !ip address 100.64.0.101/32
  ip address 100.64.0.101/32
  ip router ospf UNDERLAY area 0.0.0.0

! Define loopback interface for associating with local VTEP
interface loopback1
  description VTEP loopback interface
  !ip address 100.64.1.101/32
  ip address 100.64.1.101/32
  ip router ospf UNDERLAY area 0.0.0.0

! Configure OSPF as the underlay network
router ospf UNDERLAY
  router-id 100.64.0.101

! change ECMP hash rotate value from default 32 to 40 for better
! router port utilization for upto parallel flows via the 8
! available router ports
ip load-sharing address source-destination port source-
destination rotate 40

! - Create BGP route reflector to exchange routes across VTEPs
! Define max config 8 neighbor spines using their loopback IPs
! - BGP peers are located in an autonomous system (AS) that uses
! 4-byte AS numbers. Cisco recommends to pick a high value such
! as 65502 to avoid conflict with future bgp peers.
! - Create a template 'BasePolicy' that defines a peer policy
! template to define attributes for a particular address family.
router bgp 65502
```

```
router-id 100.64.0.101
log-neighbor-changes
```

- e. Verify the IP address replacements in each leaf switch configuration file.

For example, for the first rack check for IP addresses 100.64.0.101 and 100.64.1.101 in the lower leaf switch configuration file (R1LL), and for IP addresses 100.64.0.102 and 100.64.1.102 in the upper leaf switch configuration file (R1UL):

```
$ grep 100.64. roce_leaf_switch_multi_R1LL.cfg | grep -v neighbor |
grep -v '!'
```

```
ip address 100.64.0.101/32
ip address 100.64.1.101/32
router-id 100.64.0.101
router-id 100.64.0.101
```

```
$ grep 100.64. roce_leaf_switch_multi_R1UL.cfg | grep -v neighbor |
grep -v '!'
```

```
ip address 100.64.0.102/32
ip address 100.64.1.102/32
router-id 100.64.0.102
router-id 100.64.0.102
```

- f. Apply the updated multi-rack configuration files to each corresponding leaf switch.
- i. Log in to each leaf switch, and remove the existing configuration file. For example:

```
rack1sw-rocea0# delete bootflash:roce_leaf_switch.cfg
Do you want to delete "/roce_leaf_switch.cfg" ? (yes/no/abort) [y]
y
```

```
rack1sw-rocea0# delete bootflash:roce_leaf_switch_multi.cfg
No such file or directory
```

```
rack1sw-roceb0# delete bootflash:roce_leaf_switch.cfg
Do you want to delete "/roce_leaf_switch.cfg" ? (yes/no/abort) [y]
y
```

```
rack1sw-roceb0# delete bootflash:roce_leaf_switch_multi.cfg
No such file or directory
```

- ii. Log in to the server that contains the modified configuration files, and copy each file to its corresponding leaf switch. For example:

```
# scp roce_leaf_switch_multi_R1LL.cfg admin@R1LL_IP_Address:/
# scp roce_leaf_switch_multi_R1UL.cfg admin@R1UL_IP_Address:/
# scp roce_leaf_switch_multi_R2LL.cfg admin@R2LL_IP_Address:/
# scp roce_leaf_switch_multi_R2UL.cfg admin@R2UL_IP_Address:/
```


- iii. Log in to each leaf switch and verify that the modified files were copied successfully. For example:

```
rack1sw-rocea0# dir bootflash:roce_leaf_switch_multi_R1LL.cfg
    171387 Nov 20 14:41:52 2019
roce_leaf_switch_multi_R1LL.cfg
Usage for bootflash://sup-local
2583580672 bytes used
114139488256 bytes free
116723068928 bytes total
```

```
rack1sw-roceb0# dir bootflash:roce_leaf_switch_multi_R1UL.cfg
    171387 Nov 20 21:41:50 2019
roce_leaf_switch_multi_R1UL.cfg
Usage for bootflash://sup-local
2579836928 bytes used
114143232000 bytes free
116723068928 bytes total
```

```
rack2sw-rocea0# dir bootflash:roce_leaf_switch_multi_R2LL.cfg
    171387 Nov 20 14:41:52 2019
roce_leaf_switch_multi_R2LL.cfg
Usage for bootflash://sup-local
2583580672 bytes used
114139488256 bytes free
116723068928 bytes total
```

```
rack2sw-roceb0# dir bootflash:roce_leaf_switch_multi_R2UL.cfg
    171387 Nov 20 21:41:50 2019
roce_leaf_switch_multi_R2UL.cfg
Usage for bootflash://sup-local
2579836928 bytes used
114143232000 bytes free
116723068928 bytes total
```

- iv. Copy the modified configuration file into flash. For example:

```
rack1sw-rocea0# run-script
bootflash:roce_leaf_switch_multi_R1LL.cfg | grep 'none'
rack1sw-rocea0# copy running-config startup-config
```

```
rack1sw-roceb0# run-script
bootflash:roce_leaf_switch_multi_R1UL.cfg | grep 'none'
rack1sw-roceb0# copy running-config startup-config
```

```
rack2sw-rocea0# run-script
bootflash:roce_leaf_switch_multi_R2LL.cfg | grep 'none'
rack2sw-rocea0# copy running-config startup-config
```

```
rack2sw-roceb0# run-script
bootflash:roce_leaf_switch_multi_R2UL.cfg | grep 'none'
rack2sw-roceb0# copy running-config startup-config
```

 **Note:**

The `run-script` command for a leaf switch can take approximately 6 minutes to complete.

7. Use `patchmgr` to verify the configuration of the RDMA Network Fabric switches against the golden configuration files.
 - a. Log in to the server that has downloaded the RDMA Network Fabric patch ZIP file (from Step 5.a).
 - b. Change directories to the location of the `patchmgr` utility.
 - c. Create a file that contains the host name or IP address of the leaf and spine switches on both racks.

For example, create a file name `switches.lst`. The file must contain the host name or IP address for the 2 spine switches and 4 leaf switches, with each switch on a new line.

- d. Run `patchmgr` with the `--verify_config` option.

In the following command, `switches.lst` is a file that contains the switches to be queried, one per line.

```
$ ./patchmgr --roceswitches switches.lst --verify-config
```

```
2019-11-20 14:12:27 -0800 :Working: Initiate config verify on RoCE
switches from . Expect up to 6 minutes for each switch
```

```
2019-11-20 14:12:30 -0800 1 of 15 :Verifying config on switch rack1sw-
rocea0
```

```
2019-11-20 14:12:30 -0800: [INFO ] Dumping current running config
```

```

locally as file: /tmp/run.rack1sw-rocea0.cfg
2019-11-20 14:12:33 -0800: [SUCCESS ] Backed up switch config
successfully
2019-11-20 14:12:33 -0800: [INFO ] Validating running config
against template [1/3]: /tmp/patch_switch_19.3.1.0.0.191018/
roce_switch_templates/roce_leaf_switch.cfg
2019-11-20 14:12:33 -0800: [INFO ] Validating running config
against template [2/3]: /tmp/patch_switch_19.3.1.0.0.191018/
roce_switch_templates/roce_leaf_switch_multi.cfg
2019-11-20 14:12:33 -0800: [INFO ] Config matches template: /tmp/
patch_switch_19.3.1.0.0.191018/roce_switch_templates/
roce_leaf_switch_multi.cfg
2019-11-20 14:12:33 -0800: [SUCCESS ] Config validation
successful!

2019-11-20 14:12:33 -0800 2 of 15 :Verifying config on switch
rack1sw-roceb0
...

```

8. Perform the physical cabling of the switches.
 - a. In Rack 2, remove the eight existing inter-switch connections between the two leaf switches, R2UL and R2LL.
 - b. In Rack 2, cable each leaf switch using the tables in [Two-Rack Cabling for X8M Racks](#).
 - c. In Rack 1, remove the eight existing inter-switch connections between the two leaf switches, R1UL and R1LL.
 - d. In Rack 1, cable each leaf switch using the tables in [Two-Rack Cabling for X8M Racks](#).
9. Confirm each switch is available and connected.

For each of the 6 switches, confirm the output from the `show interface status` command shows `connected` and `100G`. In the following examples, the leaf switches are ports `Eth1/4` to `Eth1/7`, and `Eth1/30` to `Eth1/33`. The spine switches are ports `Eth1/5` to `Eth1/20`.

When run from a spine switch, the output should be similar to the following:

```

rack1sw-roces0# show interface status
-----
-----
Port          Name          Status   Vlan    Duplex
Speed  Type
-----
-----
mgmt0         --           connected routed   full
1000         --
-----
-----
Port          Name          Status   Vlan    Duplex
Speed  Type
-----
-----

```

```

...
Eth1/5      RouterPort5      connected routed    full    100G
QSFP-100G-CR4
Eth1/6      RouterPort6      connected routed    full    100G
QSFP-100G-SR4
Eth1/7      RouterPort7      connected routed    full    100G
QSFP-100G-CR4
Eth1/8      RouterPort8      connected routed    full    100G
QSFP-100G-SR4
Eth1/9      RouterPort9      connected routed    full    100G
QSFP-100G-CR4
Eth1/10     RouterPort10     connected routed    full    100G
QSFP-100G-SR4
Eth1/11     RouterPort11     connected routed    full    100G
QSFP-100G-CR4
Eth1/12     RouterPort12     connected routed    full    100G
QSFP-100G-SR4
Eth1/13     RouterPort13     connected routed    full    100G
QSFP-100G-CR4
Eth1/14     RouterPort14     connected routed    full    100G
QSFP-100G-SR4
Eth1/15     RouterPort15     connected routed    full    100G
QSFP-100G-CR4
Eth1/16     RouterPort16     connected routed    full    100G
QSFP-100G-SR4
Eth1/17     RouterPort17     connected routed    full    100G
QSFP-100G-CR4
Eth1/18     RouterPort18     connected routed    full    100G
QSFP-100G-SR4
Eth1/19     RouterPort19     connected routed    full    100G
QSFP-100G-CR4
Eth1/20     RouterPort20     connected routed    full    100G
QSFP-100G-SR4
Eth1/21     RouterPort21     xcvrAbsen    routed    full    100G
--
...

```

When run from a leaf switch, the output should be similar to the following:

```

rack1sw-rocea0# show interface status
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
mgmt0         --           connected  routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
...
Eth1/4      RouterPort1      connected routed    full    100G
QSFP-100G-CR4

```

```

Eth1/5      RouterPort2      connected routed    full
100G QSFP-100G-CR4
Eth1/6      RouterPort3      connected routed    full
100G QSFP-100G-CR4
Eth1/7      RouterPort4      connected routed    full
100G QSFP-100G-CR4
Eth1/8      celadm14         connected 3888    full
100G QSFP-100G-CR4
...
Eth1/29     celadm01         connected 3888    full
100G QSFP-100G-CR4
Eth1/30     RouterPort5      connected routed    full
100G QSFP-100G-SR4
Eth1/31     RouterPort6      connected routed    full
100G QSFP-100G-SR4
Eth1/32     RouterPort7      connected routed    full
100G QSFP-100G-SR4
Eth1/33     RouterPort8      connected routed    full
100G QSFP-100G-SR4
...
  
```

10. Check the neighbor discovery for every switch in racks R1 and R2.

Log in to each switch and use the `show lldp neighbors` command. Make sure that all switches are visible and check the switch ports assignment (leaf switches: ports Eth1/4 - Eth1/7, Eth1/30 - Eth1/33; spine switches: ports Eth1/5 - Eth1/20) against the tables in [Two-Rack Cabling for X8M Racks](#).

A spine switch should see the two leaf switches in each rack, but not the other spine switch. The output for a spine switch should be similar to the following:

 **Note:**

The interfaces output in the `Port ID` column are different for each switch based on the applicable cabling tables.

```

rack1sw-roces0# show lldp neighbors
...
Device ID           Local Intf           Hold-time  Capability  Port ID
rack1-adm0          mgmt0                120        BR
Ethernet1/47
rack1sw-roceb0      Eth1/5             120        BR          Ethernet1/5
rack2sw-roceb0      Eth1/6             120        BR          Ethernet1/5
rack1sw-roceb0      Eth1/7             120        BR          Ethernet1/7
rack2sw-roceb0      Eth1/8             120        BR          Ethernet1/7
rack1sw-roceb0      Eth1/9             120        BR          Ethernet1/4
rack2sw-roceb0      Eth1/10            120        BR          Ethernet1/4
rack1sw-roceb0      Eth1/11            120        BR          Ethernet1/6
rack2sw-roceb0      Eth1/12            120        BR          Ethernet1/6
rack1sw-rocea0      Eth1/13            120        BR          Ethernet1/5
rack2sw-rocea0      Eth1/14            120        BR          Ethernet1/5
rack1sw-rocea0      Eth1/15            120        BR          Ethernet1/7
rack2sw-rocea0      Eth1/16            120        BR          Ethernet1/7
  
```

```

rack1sw-rocea0    Eth1/17    120    BR    Ethernet1/4
rack2sw-rocea0    Eth1/18    120    BR    Ethernet1/4
rack1sw-rocea0    Eth1/19    120    BR    Ethernet1/6
rack2sw-rocea0    Eth1/20    120    BR    Ethernet1/6
Total entries displayed: 17

```

Each leaf switch should see the two spine switches, but not the other leaf switches. The output for a leaf switch should be similar to the following:

 **Note:**

The interfaces output in the Port ID column are different for each switch based on the applicable cabling tables.

```

rack1sw-rocea0# show lldp neighbors
...
Device ID          Local Intf      Hold-time  Capability  Port ID
switch            mgmt0          120       BR         Ethernet1/46
rack1sw-roces0    Eth1/4         120       BR         Ethernet1/17
rack1sw-roces0    Eth1/5         120       BR         Ethernet1/13
rack1sw-roces0    Eth1/6         120       BR         Ethernet1/19
rack1sw-roces0    Eth1/7         120       BR         Ethernet1/15
rack2sw-roces0    Eth1/30        120       BR         Ethernet1/17
rack2sw-roces0    Eth1/31        120       BR         Ethernet1/13
rack2sw-roces0    Eth1/32        120       BR         Ethernet1/19
rack2sw-roces0    Eth1/33        120       BR         Ethernet1/15
rocetoi-ext-sw    Eth1/36        120       BR         Ethernet1/49
Total entries displayed: 10

```

11. Power on all servers in racks R1 and R2.
12. For each rack, confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

Refer to My Oracle Support Doc ID 2587717.1 for download and usage instructions.

Check the output of the `verify_roce_cables.py` script against the tables in [Two-Rack Cabling for X8M Racks](#). Also, check that output in the `CABLE OK?` columns contains the OK status.

When running the script, two input files are used, one for nodes and one for switches. Each file should contain the servers or switches on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

The following output is a partial example of the command results:

```

# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER)  LEAF-1 (rack1sw-rocea0)      : CABLE OK?  LEAF-2 (rack1sw-roceb0)      : CABLE OK?
-----
Eth1/4 (ISL peer switch)     : rack1sw-roces0 Ethernet1/17 : OK          rack1sw-roces0
Ethernet1/9 : OK
Eth1/5 (ISL peer switch)     : rack1sw-roces0 Ethernet1/13 : OK          rack1sw-roces0

```

```

Ethernet1/5 : OK
Eth1/6 (ISL peer switch) : rack1sw-roces0 Ethernet1/19 : OK      rack1sw-
roces0 Ethernet1/11: OK
Eth1/7 (ISL peer switch) : rack1sw-roces0 Ethernet1/15 : OK      rack1sw-
roces0 Ethernet1/7 : OK
Eth1/12 (celadm10)      : rack1celadm10 port-1      : OK
rack1celadm10 port-2    : OK
Eth1/13 (celadm09)      : rack1celadm09 port-1      : OK
rack1celadm09 port-2    : OK
Eth1/14 (celadm08)      : rack1celadm08 port-1      : OK
rack1celadm08 port-2    : OK
...
Eth1/15 (adm08)         : rack1dbadm08 port-1      : OK      rack1dbadm08
port-2      : OK
Eth1/16 (adm07)         : rack1dbadm07 port-1      : OK      rack1dbadm07
port-2      : OK
Eth1/17 (adm06)         : rack1dbadm06 port-1      : OK      rack1dbadm06
port-2      : OK
...
Eth1/30 (ISL peer switch) : rack2sw-roces0 Ethernet1/17 : OK      rack2sw-
roces0 Ethernet1/9 : OK
Eth1/31 (ISL peer switch) : rack2sw-roces0 Ethernet1/13 : OK      rack2sw-
roces0 Ethernet1/5 : OK
Eth1/32 (ISL peer switch) : rack2sw-roces0 Ethernet1/19 : OK      rack2sw-
roces0 Ethernet1/11: OK
Eth1/33 (ISL peer switch) : rack2sw-roces0 Ethernet1/15 : OK      rack2sw-
roces0 Ethernet1/7 : OK

# ./verify_roce_cables.py -n nodes.rack2 -s switches.rack2
SWITCH PORT (EXPECTED PEER) LEAF-1 (rack2sw-rocea0) : CABLE OK? LEAF-2
(rack2sw-roceb0) : CABLE OK?
----- : -----
: -----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/18 : OK      rack1sw-
roces0 Ethernet1/10: OK
...

```

13. Verify the RoCE Network Fabric operation by using the `infinicheck` command.

Use the following recommended command sequence. In each command, `hosts.lst` is the name of an input file that contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses, and `cells.lst` is the name of an input file that contains a list of RoCE Network Fabric IP addresses for the storage servers.

- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -
c cells.lst -z
```

- Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -s
```

- Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -b
```

```
INFINICHECK
      [Network Connectivity, Configuration and Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS].... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS].... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS].... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs and
CELLS
[SUCCESS].... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value : 2300
[SUCCESS].... RoCE interface MTU settings look good
      Verifying switch advertised DSCP on all DBs and CELLS ports ( )
[SUCCESS].... Advertised DSCP settings from RoCE switch looks good
      ##### CONNECTIVITY TESTS #####
      [COMPUTE NODES -> STORAGE CELLS]
      (60 seconds approx.)
      (Will walk through QoS values: 0-6) [SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
      [COMPUTE NODES -> COMPUTE NODES]
...

```

14. After cabling the racks together, proceed to [Configuring the New Hardware](#) to finish the configuration of the new rack.

Related Topics

- [Exadata Database Machine and Exadata Storage Server Supported Versions \(My Oracle Support Doc ID 888828.1\)](#)
- [Verify RoCE Cabling on Oracle Exadata Database Machine X8M-2 and X8M-8 Servers \(My Oracle Support Doc ID 2587717.1\)](#)

2.3.2.3 Cabling Two InfiniBand Network Fabric Racks Together

Use this procedure to cable together two racks with InfiniBand Network Fabric.

This procedure assumes that the racks are adjacent to each other.

In the procedure, the existing rack is R1, and the new rack is R2.

1. Set the priority of the current, active Subnet Manager Master to 10 on the spine switch, as follows:
 - a. Log in to any RDMA Network Fabric switch on the active system.
 - b. Use the `getmaster` command to determine that the Subnet Manager Master is running on the spine switch. If it is not, then follow the procedure [Setting the Subnet Manager Master on Oracle Exadata Database Machine Full Rack and Oracle Exadata Database Machine Half Rack](#) in *Oracle Exadata Database Machine Installation and Configuration Guide*.
 - c. Log in to the spine switch.
 - d. Use the `disablesm` command to stop Subnet Manager.
 - e. Use the `setsmpriority 10` command to set the priority to 10.
 - f. Use the `enablesm` command to restart Subnet Manager.
 - g. Repeat step 1.b to ensure the Subnet Manager Master is running on the spine switch.
2. Ensure the new rack is near the existing rack. The RDMA Network Fabric cables must be able to reach the servers in each rack.
3. Completely shut down the new rack (R2).
4. Cable the two leaf switches R2 IB2 and R2 IB3 in the new rack according to [Two-Rack Cabling with InfiniBand Network Fabric](#). Note that you need to first remove the seven existing inter-switch connections between each leaf switch, as well as the two connections between the leaf switches and the spine switch in the new rack R2, not in the existing rack R1.
5. Verify both RDMA Network Fabric interfaces are up on all database nodes and storage cells. You can do this by running the `ibstat` command on each node and verifying both interfaces are up.
6. Power off leaf switch R1 IB2. This causes all the database servers and Exadata Storage Servers to fail over their RDMA Network Fabric traffic to R1 IB3.
7. Disconnect all seven inter-switch links between R1 IB2 and R1 IB3, as well as the one connection between R1 IB2 and the spine switch R1 IB1.
8. Cable leaf switch R1 IB2 according to [Two-Rack Cabling with InfiniBand Network Fabric](#).
9. Power on leaf switch R1 IB2.
10. Wait for three minutes for R1 IB2 to become completely operational.

To check the switch, log in to the switch and run the `ibswitches` command. The output should show three switches, R1 IB1, R1 IB2, and R1 IB3.
11. Verify both RDMA Network Fabric interfaces are up on all database nodes and storage cells. You can do this by running the `ibstat` command on each node and verifying both interfaces are up.

12. Power off leaf switch R1 IB3. This causes all the database servers and storage servers to fail over their RDMA Network Fabric traffic to R1 IB2.
13. Disconnect the one connection between R1 IB3 and the spine switch R1 IB1.
14. Cable leaf switch R1 IB3 according to [Two-Rack Cabling with InfiniBand Network Fabric](#).
15. Power on leaf switch R1 IB3.
16. Wait for three minutes for R1 IB3 to become completely operational.
To check the switch, log in to the switch and run the `ibswitches` command. The output should show three switches, R1 IB1, R1 IB2, and R1 IB3.
17. Power on all the InfiniBand switches in R2.
18. Wait for three minutes for the switches to become completely operational.
To check the switch, log in to the switch and run the `ibswitches` command. The output should show six switches, R1 IB1, R1 IB2, R1 IB3, R2 IB1, R2 IB2, and R2 IB3.
19. Ensure the Subnet Manager Master is running on R1 IB1 by running the `getmaster` command from any switch.
20. Power on all servers in R2.
21. Log in to spine switch R1 IB1, and lower its priority to 8 as follows:
 - a. Use the `disableesm` command to stop Subnet Manager.
 - b. Use the `setsmpriority 8` command to set the priority to 8.
 - c. Use the `enableesm` command to restart Subnet Manager.
22. Ensure Subnet Manager Master is running on one of the spine switches.

After cabling the racks together, proceed to [Configuring the New Hardware](#) to configure the racks.

2.3.3 Cabling Several Racks Together

By using the following procedures, you can add another rack to an existing multi-rack system.

The procedures assume that the racks are adjacent to each other. The existing racks are R1, R2, ... Rn, and the new rack is Rn+1. For example, if you have four racks and you are adding a fifth rack, the existing racks are designated R1, R2, R3, and R4 and the new rack is designated R5.

Starting with Exadata Database Machine X9M models, you can cable up to 12 racks together without additional switches. For prior models (up to X8M), you can cable up to 18 racks together without additional switches.

The procedures differ for systems that use RoCE Network Fabric (X8M and later) and InfiniBand Network Fabric (X8 and earlier).

- [Cabling Several RoCE Network Fabric Racks Together using Oracle Exadata System Software Release 20.1.0 or Later](#)
Use this procedure to add another rack to an existing multi-rack system with RoCE Network Fabric using Oracle Exadata System Software Release 20.1.0 or later.
- [Cabling Several RoCE Network Fabric Racks Together using Oracle Exadata System Software Release 19.3 or Earlier](#)
Use this procedure to add another rack to an existing multi-rack system with RoCE Network Fabric using Oracle Exadata System Software Release 19.3 or earlier.

- [Cabling Several InfiniBand Network Fabric Racks Together](#)
Use this procedure to add another rack to an existing multi-rack system with InfiniBand Network Fabric.

2.3.3.1 Cabling Several RoCE Network Fabric Racks Together using Oracle Exadata System Software Release 20.1.0 or Later

Use this procedure to add another rack to an existing multi-rack system with RoCE Network Fabric using Oracle Exadata System Software Release 20.1.0 or later.

This procedure is for systems with RoCE Network Fabric (X8M or later).

In this procedure, the existing racks are R1, R2, ..., Rn, and the new rack is Rn+1. In the following steps, these example switch names are used:

- rack5sw-roces0: Rack 5 Spine switch (SS)
- rack5sw-rocea0: Rack 5 Lower Leaf switch (R5LL)
- rack5sw-roceb0: Rack 5 Upper Leaf switch (R5UL)



Note:

Cabling three or more racks together requires no downtime for the existing racks R1, R2, ..., Rn. Only the new rack, Rn+1, is powered down

Use the applicable cabling tables depending on your system:

- [Multi-Rack Cabling Tables for Oracle Exadata Rack X9M](#)
- [Multi-Rack Cabling Tables for Oracle Exadata Rack X8M](#)

1. Ensure the new rack is near the existing racks R1, R2, ..., Rn.

The RDMA Network Fabric cables must be able to reach the servers in each rack.

2. Ensure you have a backup of the current switch configuration for each switch in the existing racks and the new rack.

For each switch, complete the steps in the Oracle Exadata Database Machine Maintenance Guide, section Backing Up Settings on the RoCE Network Fabric Switch.

3. Shut down all servers in the new rack Rn+1.

Refer to Powering Off Oracle Exadata Rack. The switches must remain online and available.

4. Apply the golden configuration settings on the RoCE Network Fabric switches in the new rack Rn+1.

Use the procedure described in Applying Golden Configuration Settings on RoCE Network Fabric Switches, in *Oracle Exadata Database Machine Maintenance Guide*.

5. Perform the physical cabling of the switches in the new rack Rn+1.

▲ Caution:

Cabling within a live network must be done carefully in order to avoid potentially serious disruptions.

- a. Remove the eight existing inter-switch connections between each leaf switch in the new rack R_{n+1} (ports 4, 5, 6, 7 and 30, 31, 32, 33).
- b. Cable the leaf switches in the new rack according to the applicable cabling table.
 For example, if you are adding a 5th rack and rack R_{n+1} is R5, then use "Table 5-14 Leaf Switch Connections for the Fifth Rack in a Five-Rack System".
6. Add the new rack to the switches in the existing racks (R1 to R_n).
 - a. For an existing rack (Rx), cable the lower leaf switch RxLL according to the applicable cabling table.
 - b. For the same rack, cable the upper leaf switch RxUL according to the applicable cabling table.
 - c. Repeat these steps for each existing rack, R1 to R_n .
7. Confirm each switch is available and connected.

For each switch in racks R1, R2, ..., R_n , R_{n+1} , confirm the output for the switch `show interface status` command shows `connected` and `100G`. In the following example, the leaf switches are ports Eth1/4 to Eth1/7, and Eth1/30 to Eth1/33. The spine switches are ports Eth1/5 to Eth1/20.

When run from a spine switch, the output should be similar to the following:

```
rack1sw-roces0# show interface status
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
mgmt0         --           connected  routed   full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
...
Eth1/5       RouterPort5   connected  routed   full    100G
QSF-100G-CR4
Eth1/6       RouterPort6   connected  routed   full    100G
QSF-100G-SR4
Eth1/7       RouterPort7   connected  routed   full    100G
QSF-100G-CR4
Eth1/8       RouterPort8   connected  routed   full    100G
QSF-100G-SR4
Eth1/9       RouterPort9   connected  routed   full    100G
QSF-100G-CR4
Eth1/10      RouterPort10  connected  routed   full    100G
QSF-100G-SR4
Eth1/11      RouterPort11  connected  routed   full    100G
```

```

QSFPl-100G-CR4
Eth1/12      RouterPort12      connected routed    full
100G QSFPl-100G-SR4
Eth1/13      RouterPort13      connected routed    full
100G QSFPl-100G-CR4
Eth1/14      RouterPort14      connected routed    full
100G QSFPl-100G-SR4
Eth1/15      RouterPort15      connected routed    full
100G QSFPl-100G-CR4
Eth1/16      RouterPort16      connected routed    full
100G QSFPl-100G-SR4
Eth1/17      RouterPort17      connected routed    full
100G QSFPl-100G-CR4
Eth1/18      RouterPort18      connected routed    full
100G QSFPl-100G-SR4
Eth1/19      RouterPort19      connected routed    full
100G QSFPl-100G-CR4
Eth1/20      RouterPort20      connected routed    full
100G QSFPl-100G-SR4
Eth1/21      RouterPort21      xcvrAbsen    routed    full
100G      --
...
  
```

When run from a leaf switch, the output should be similar to the following:

```

rack1sw-rocea0# show interface status
-----
-----
Port      Name      Status  Vlan  Duplex
Speed  Type
-----
-----
mgmt0      --      connected routed    full
1000      --
-----
-----
Port      Name      Status  Vlan  Duplex
Speed  Type
-----
-----
...
Eth1/4      RouterPort1      connected routed    full
100G QSFPl-100G-CR4
Eth1/5      RouterPort2      connected routed    full
100G QSFPl-100G-CR4
Eth1/6      RouterPort3      connected routed    full
100G QSFPl-100G-CR4
Eth1/7      RouterPort4      connected routed    full
100G QSFPl-100G-CR4
Eth1/8      celadm14      connected 3888    full
100G      QSFPl-100G-CR4
...
Eth1/29      celadm01      connected 3888    full
100G      QSFPl-100G-CR4
  
```

```

Eth1/30      RouterPort5      connected routed    full    100G
QSFP-100G-SR4
Eth1/31      RouterPort6      connected routed    full    100G
QSFP-100G-SR4
Eth1/32      RouterPort7      connected routed    full    100G
QSFP-100G-SR4
Eth1/33      RouterPort8      connected routed    full    100G
QSFP-100G-SR4
...

```

8. Check the neighbor discovery for every switch in racks R1, R2, ..., Rn, Rn+1.

Log in to each switch and use the `show lldp neighbors` command. Make sure that all switches are visible and check the switch ports assignment (leaf switches: ports Eth1/4 - Eth1/7, Eth1/30 - Eth1/33; spine switches: ports Eth1/5 - Eth1/20) against the applicable cabling tables.

Each spine switch should see all the leaf switches in each rack, but not the other spine switches. The output for a spine switch should be similar to the following:

 **Note:**

The interfaces in the rightmost output column (for example, Ethernet1/5) are different for each switch based on the applicable cabling tables.

```

rack1sw-roces0# show lldp neighbors | grep roce
rack1sw-roceb0 Eth1/5 120 BR Ethernet1/5
rack2sw-roceb0 Eth1/6 120 BR Ethernet1/5
rack1sw-roceb0 Eth1/7 120 BR Ethernet1/7
rack2sw-roceb0 Eth1/8 120 BR Ethernet1/7
rack1sw-roceb0 Eth1/9 120 BR Ethernet1/4
rack2sw-roceb0 Eth1/10 120 BR Ethernet1/4
rack3sw-roceb0 Eth1/11 120 BR Ethernet1/5
rack3sw-roceb0 Eth1/12 120 BR Ethernet1/7
rack1sw-rocea0 Eth1/13 120 BR Ethernet1/5
rack2sw-rocea0 Eth1/14 120 BR Ethernet1/5
rack1sw-rocea0 Eth1/15 120 BR Ethernet1/7
rack2sw-rocea0 Eth1/16 120 BR Ethernet1/7
rack3sw-rocea0 Eth1/17 120 BR Ethernet1/5
rack2sw-rocea0 Eth1/18 120 BR Ethernet1/4
rack3sw-rocea0 Eth1/19 120 BR Ethernet1/7
rack3sw-rocea0 Eth1/20 120 BR Ethernet1/4

```

Each leaf switch should see the spine switch in every rack, but not the other leaf switches. The output for a leaf switch should be similar to the following:

 **Note:**

The interfaces in the rightmost output column (for example, Ethernet1/13) are different for each switch based on the applicable cabling tables.

```
rack1sw-rocea0# show lldp neighbors | grep roce
rack3sw-roces0 Eth1/4 120 BR Ethernet1/13
rack1sw-roces0 Eth1/5 120 BR Ethernet1/13
rack3sw-roces0 Eth1/6 120 BR Ethernet1/15
rack1sw-roces0 Eth1/7 120 BR Ethernet1/15
rack2sw-roces0 Eth1/30 120 BR Ethernet1/17
rack2sw-roces0 Eth1/31 120 BR Ethernet1/13
rack3sw-roces0 Eth1/32 120 BR Ethernet1/17
rack2sw-roces0 Eth1/33 120 BR Ethernet1/15
```

9. Power on all the servers in the new rack, $Rn+1$.
10. For each rack, confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

Refer to My Oracle Support Doc ID 2587717.1 for download and usage instructions.

Check the output of the `verify_roce_cables.py` script against the applicable cabling tables. Also, check that output in the `CABLE OK?` columns contains the `OK` status.

When running the script, two input files are used, one for nodes and one for switches. Each file should contain the servers or switches on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

The following output is a partial example of the command results:

```
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER) LEAF-1 (rack1sw-rocea0) : CABLE OK? LEAF-2
(rack1sw-roceb0) : CABLE OK?
-----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/17 : OK rack1sw-
roces0 Ethernet1/9 : OK
Eth1/5 (ISL peer switch) : rack1sw-roces0 Ethernet1/13 : OK rack1sw-
roces0 Ethernet1/5 : OK
Eth1/6 (ISL peer switch) : rack1sw-roces0 Ethernet1/19 : OK rack1sw-
roces0 Ethernet1/11: OK
Eth1/7 (ISL peer switch) : rack1sw-roces0 Ethernet1/15 : OK rack1sw-
roces0 Ethernet1/7 : OK
Eth1/12 (celadm10) : rack1celadm10 port-1 : OK
rack1celadm10 port-2 : OK
Eth1/13 (celadm09) : rack1celadm09 port-1 : OK
rack1celadm09 port-2 : OK
Eth1/14 (celadm08) : rack1celadm08 port-1 : OK
rack1celadm08 port-2 : OK
...
Eth1/15 (adm08) : rack1dbadm08 port-1 : OK rack1dbadm08
```

```

port-2          : OK
Eth1/16 (adm07) : rack1dbadm07 port-1          : OK          rack1dbadm07
port-2          : OK
Eth1/17 (adm06) : rack1dbadm06 port-1          : OK          rack1dbadm06
port-2          : OK
...
Eth1/30 (ISL peer switch) : rack2sw-roces0 Ethernet1/17 : OK          rack2sw-roces0
Ethernet1/9 : OK
Eth1/31 (ISL peer switch) : rack2sw-roces0 Ethernet1/13 : OK          rack2sw-roces0
Ethernet1/5 : OK
Eth1/32 (ISL peer switch) : rack2sw-roces0 Ethernet1/19 : OK          rack2sw-roces0
Ethernet1/11: OK
Eth1/33 (ISL peer switch) : rack2sw-roces0 Ethernet1/15 : OK          rack2sw-roces0
Ethernet1/7 : OK

```

11. Verify the RoCE Network Fabric operation by using the `infinicheck` command.

Use the following recommended command sequence. In each command, `hosts.lst` is the name of an input file that contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses, and `cells.lst` is the name of an input file that contains a list of RoCE Network Fabric IP addresses for the storage servers.

- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -z
```

- Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -s
```

- Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -c
cells.lst -b
```

```

INFINICHECK
      [Network Connectivity, Configuration and Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all CELLS.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS]... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and CELLS
[SUCCESS]... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS]... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs and

```



```

CELLs
[SUCCESS].... RoCE PFC and DSCP settings look good
    Checking for RoCE interface MTU settings. Expected value :
2300
[SUCCESS].... RoCE interface MTU settings look good
    Verifying switch advertised DSCP on all DBs and CELLs ports
( )
[SUCCESS].... Advertised DSCP settings from RoCE switch looks
good
    ##### CONNECTIVITY TESTS #####
    [COMPUTE NODES -> STORAGE CELLS]
    (60 seconds approx.)
    (Will walk through QoS values: 0-6)
[SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
    [COMPUTE NODES -> COMPUTE NODES]
...
  
```

12. After cabling the racks together, proceed to [Configuring the New Hardware](#) to finish the configuration of the new rack.

Related Topics

- [Exadata Database Machine and Exadata Storage Server Supported Versions \(My Oracle Support Doc ID 888828.1\)](#)

2.3.3.2 Cabling Several RoCE Network Fabric Racks Together using Oracle Exadata System Software Release 19.3 or Earlier

Use this procedure to add another rack to an existing multi-rack system with RoCE Network Fabric using Oracle Exadata System Software Release 19.3 or earlier.

This procedure is for systems with RoCE Network Fabric (X8M or later).

In this procedure, the existing racks are R1, R2, ..., Rn, and the new rack is Rn+1. In the following steps, these example switch names are used:

- rack5sw-roces0: Rack 5 Spine switch (SS)
- rack5sw-rocea0: Rack 5 Lower Leaf switch (R5LL)
- rack5sw-roceb0: Rack 5 Upper Leaf switch (R5UL)

Note:

Cabling three or more racks together requires no downtime for the existing racks R1, R2, ..., Rn. Only the new rack, Rn+1, is powered down

1. Ensure the new rack is near the existing racks R1, R2, ..., Rn.
 The RDMA Network Fabric cables must be able to reach the servers in each rack.
2. Ensure you have a backup of the current switch configuration for each switch in the existing racks and the new rack.

For each switch, complete the steps in the Oracle Exadata Database Machine Maintenance Guide, section Backing Up Settings on the RoCE Network Fabric Switch.

3. Shut down all servers in the new rack R_{n+1} .

Refer to Powering Off Oracle Exadata Rack. The switches must remain online and available.

4. Apply the multi-rack spine switch configuration to the spine switch in the new rack R_{n+1} :

- a. Log in to the server that has downloaded the latest RDMA Network Fabric patch ZIP file.

To find the available RDMA Network Fabric patches, search for '**RDMA network switch**' in My Oracle Support document 888828.1. Download and use the latest patch for your Oracle Exadata System Software release.

- b. Unzip the RDMA Network Fabric patch ZIP file and change directories to the location of the `patchmgr` utility.

- c. Make a copy of the golden configuration file for the new spine switch.

Run these commands from patch directory, where $n+1$ is the number of the new rack:

```
# cp roce_switch_templates/roce_spine_switch_multi.cfg
roce_spine_switch_multi_Rn+1SS.cfg
```

- d. Edit the copy of the spine switch configuration file.

Using a text editor, replace the three occurrences of `%SPINE_LOOPBACK_IP0%` with the correct IP address for the switch, as indicated in the table below, using the value that matches R_{n+1} for your environment.

Switch	SPINE_LOOPBACK_IP0
Rack 3 spine switch (R3SS)	100.64.0.203
Rack 4 spine switch (R4SS)	100.64.0.204
Rack 5 spine switch (R5SS)	100.64.0.205
Rack 6 spine switch (R6SS)	100.64.0.206
Rack 7 spine switch (R7SS)	100.64.0.207
Rack 8 spine switch (R8SS)	100.64.0.208

For example, if you are adding a rack to an existing 4-rack system (where $n+1=5$), then use IP address 100.64.0.205 as the `SPINE_LOOPBACK_IP0` for the spine switch in the new rack (R5SS).

```
! Define loopback interface for underlay OSPF routing
interface loopback0
  description Routing loopback interface
  !ip address 100.64.0.201/32
  ip address 100.64.0.205/32
  ip router ospf UNDERLAY area 0.0.0.0
```

```
! Configure OSPF as the underlay network
router ospf UNDERLAY
  router-id 100.64.0.205
```

```
! change ECMP hash rotate value from default 32 to 40 for better
```

```

! router port utilization for upto parallel flows via the 8
! available router ports
ip load-sharing address source-destination port source-
destination rotate 40

! Create BGP route reflector to exchange routes across VTEPs
! Use CIDR block of IPs for neighbor range
! - log-neighbor-changes: Enables the generation of logging
messages
! generated when the status of a BGP neighbor changes.
! - address-family ipv4 unicast: Enters address family
configuration
! mode and Specifies IP Version 4 unicast address prefixes.
! address
router bgp 65502
  router-id 100.64.0.205
  log-neighbor-changes

```

- e. Verify the three replacements in the spine switch configuration file.

For example, if you are adding a 5th rack, then check for IP address 100.64.0.205 in the spine switch configuration file:

```

$ grep 100.64 roce_spine_switch_multi_R5SS.cfg |grep -v
'neighbor' |grep -v '!
ip address 100.64.0.205/32
router-id 100.64.0.205
router-id 100.64.0.205

```

- f. Apply the updated multi-rack configuration file to the spine switch in the new rack R_{n+1} :
- i. Log in to the switch in the new rack R_{n+1} , and remove the existing configuration file, if it exists. For example, if you are adding a 5th rack, you would use the following command:

```

rack5sw-roces0# delete bootflash:roce_spine_switch_multi.cfg
Do you want to delete "/roce_spine_switch_multi.cfg" ?
(yes/no/abort) [y] y
rack5sw-roces0#

```

- ii. Log in to the server that contains the modified configuration file for the spine switch, and copy the file to the spine switch in the new rack. For example, if you are adding a 5th rack:

```

# scp roce_spine_switch_multi_R5SS.cfg
admin@R5SS_IP_Address:/

```

- iii. Verify the modified file was copied successfully to the spine switch. For example, if you are adding a 5th rack, log in to the spine switch on the new rack R_{n+1} again and use the following command:

```

rack5sw-roces0# dir
bootflash:roce_spine_switch_multi_R5SS.cfg
27360 Nov 20 12:12:50 2019

```

```
roce_spine_switch_multi_R5SS.cfg
Usage for bootflash://sup-local
1829572608 bytes used
114893496320 bytes free
116723068928 bytes total
```

iv. Copy the modified configuration into flash.

For example, if you are adding a 5th rack, you would use the following commands:

```
rack5sw-roces0# run-script
bootflash:roce_spine_switch_multi_R5SS.cfg | grep 'none'

rack5sw-roces0# copy running-config startup-config
```

 **Note:**

The `run-script` command for a spine switch can take approximately 2 minutes to complete.

5. Apply the multi-rack leaf switch configuration to the leaf switches in the new rack $Rn+1$:

For each leaf switch, complete the following steps, where $SW\#$ represents the values $Rn+1LL$ or $Rn+1UL$, depending on which switch you are configuring.

- a. Log in to the server that has downloaded the RDMA Network Fabric patch ZIP file (from Step 4.a) for the Oracle Exadata System Software release used by the existing racks.
- b. Change directories to the location of the `patchmgr` utility.
- c. Make a copy of the golden configuration file for each leaf switch.

You can copy either the `roce_leaf_switch_multi.cfg` file or the `roce_qinq_leaf_switch_multi.cfg` file if you want to enable Secure Fabric on the rack.

Run the following command twice from the patch directory, substituting for $SW\#$ the values $Rn+1LL$ and $Rn+1UL$.

```
# cp roce_switch_templates/roce_leaf_switch_multi.cfg
roce_leaf_switch_multi_ $SW\#$ .cfg
```

- d. Edit each copy of the leaf switch configuration file to replace the loopback IP addresses:

Using a text editor, replace the three occurrences of `%LEAF_LOOPBACK_IP0%` and one occurrence of `%LEAF_LOOPBACK_IP1%` with the correct IP addresses for the leaf switch, as indicated in the table below.

Switch	LEAF_LOOPBACK_IP0	LEAF_LOOPBACK_IP1
Rack 3 Lower Leaf switch (R3LL)	100.64.0.105	100.64.1.105
Rack 3 Upper Leaf switch (R3UL)	100.64.0.106	100.64.1.106
Rack 4 Lower Leaf switch (R4LL)	100.64.0.107	100.64.1.107
Rack 4 Upper Leaf switch (R4UL)	100.64.0.108	100.64.1.108
Rack 5 Lower Leaf switch (R5LL)	100.64.0.109	100.64.1.109
Rack 5 Upper Leaf switch (R5UL)	100.64.0.110	100.64.1.110
Rack 6 Lower Leaf switch (R6LL)	100.64.0.111	100.64.1.111
Rack 6 Upper Leaf switch (R6UL)	100.64.0.112	100.64.1.112
Rack 7 Lower Leaf switch (R7LL)	100.64.0.113	100.64.1.113
Rack 7 Upper Leaf switch (R7UL)	100.64.0.114	100.64.1.114
Rack 8 Lower Leaf switch (R8LL)	100.64.0.115	100.64.1.115
Rack 8 Upper Leaf switch (R8UL)	100.64.0.116	100.64.1.116

For example, if you are adding a 5th rack to an existing 4-rack system, then the configuration file for the lower leaf switch on rack 5 (R5LL) would look like the following:

```
! Define loopback interface for IGP protocol for VTEP
reachability
interface loopback0
  description Routing loopback interface
  !ip address 100.64.0.101/32
  ip address 100.64.0.109/32
  ip router ospf UNDERLAY area 0.0.0.0

! Define loopback interface for associating with local VTEP
interface loopback1
  description VTEP loopback interface
  !ip address 100.64.1.101/32
  ip address 100.64.1.109/32
  ip router ospf UNDERLAY area 0.0.0.0

! Configure OSPF as the underlay network
router ospf UNDERLAY
  router-id 100.64.0.109

! change ECMP hash rotate value from default 32 to 40 for better
! router port utilization for upto parallel flows via the 8
! available router ports
ip load-sharing address source-destination port source-
```

```
destination rotate 40

! - Create BGP route reflector to exchange routes across VTEPs
! Define max config 8 neighbor spines using their loopback IPs
! - BGP peers are located in an autonomous system (AS) that uses
! 4-byte AS numbers. Cisco recommends to pick a high value such
! as 65502 to avoid conflict with future bgp peers.
! - Create a template 'BasePolicy' that defines a peer policy
! template to define attributes for a particular address family.
router bgp 65502
  router-id 100.64.0.109
  log-neighbor-changes
```

- e. Verify the IP address replacements in each leaf switch configuration file.

For example, if you are adding a 5th rack, then check for IP addresses 100.64.0.109 and 100.64.1.109 in the lower leaf switch configuration file (R5LL), and for IP addresses 100.64.0.110 and 100.64.1.110 in the upper leaf switch configuration file (R5UL):

```
$ grep 100.64. roce_leaf_switch_multi_R5LL.cfg | grep -v neighbor |
grep -v '!'
```

```
ip address 100.64.0.109/32
ip address 100.64.1.109/32
router-id 100.64.0.109
router-id 100.64.0.109
```

```
$ grep 100.64. roce_leaf_switch_multi_R5UL.cfg | grep -v neighbor |
grep -v '!'
```

```
ip address 100.64.0.110/32
ip address 100.64.1.110/32
router-id 100.64.0.110
router-id 100.64.0.110
```

- f. Apply the updated multi-rack configuration files to each corresponding leaf switch in the new rack:
- i. Log in to each leaf switch, and remove the existing configuration file. For example:

```
rack5sw-rocea0# delete bootflash:roce_leaf_switch.cfg
Do you want to delete "/roce_leaf_switch.cfg" ? (yes/no/abort) [y]
y
```

```
rack5sw-rocea0# delete bootflash:roce_leaf_switch_multi.cfg
No such file or directory
```

```
rack5sw-roceb0# delete bootflash:roce_leaf_switch.cfg
Do you want to delete "/roce_leaf_switch.cfg" ? (yes/no/abort) [y]
y
```

```
rack5sw-roceb0# delete bootflash:roce_leaf_switch_multi.cfg
No such file or directory
```

- ii. Log in to the server that contains the modified configuration files, and copy each file to its corresponding leaf switch. For example:

```
# scp roce_leaf_switch_multi_R5LL.cfg admin@rack5sw-rocea0:/
User Access Verification
Password:
roce_leaf_switch_multi_R5LL.cfg 100% 167KB 487.6KB/s 00:00

# scp roce_leaf_switch_multi_R5UL.cfg admin@rack5sw-roceb0:/
User Access Verification
Password:
roce_leaf_switch_multi_R5UL.cfg
```

- iii. Log in to each leaf switch and verify that the modified files were copied successfully. For example:

```
rack5sw-rocea0# dir bootflash:roce_leaf_switch_multi_R5LL.cfg
 171387 Nov 20 14:41:52 2019
roce_leaf_switch_multi_R5LL.cfg
Usage for bootflash://sup-local
2583580672 bytes used
114139488256 bytes free
116723068928 bytes total
```

```
rack5sw-roceb0# dir bootflash:roce_leaf_switch_multi_R5UL.cfg
 171387 Nov 20 21:41:50 2019
roce_leaf_switch_multi_R5UL.cfg
Usage for bootflash://sup-local
2579836928 bytes used
114143232000 bytes free
116723068928 bytes total
```

- iv. Copy the modified configuration file into flash. For example:

```
rack5sw-rocea0# run-script
bootflash:roce_leaf_switch_multi_R5LL.cfg | grep 'none'

rack5sw-rocea0# copy running-config startup-config
```

```
rack5sw-roceb0# run-script
bootflash:roce_leaf_switch_multi_R5UL.cfg | grep 'none'

rack5sw-roceb0# copy running-config startup-config
```

 **Note:**

The `run-script` command for a leaf switch can take approximately 6 minutes to complete.

6. Use `patchmgr` to verify the configuration of the RDMA Network Fabric switches against the golden configuration files.

- a. Log in to the server that has downloaded the RDMA Network Fabric patch ZIP file (from Step 4.a).
- b. Change directories to the location of the `patchmgr` utility.
- c. Create a file that contains the host name or IP address of the leaf and spine switches on all racks.

For example, create a file named `switches.lst`. The file contains the host name or IP address for the spine switches and both leaf switches on each rack, with each switch on a new line.

- d. Run `patchmgr` with the `--verify_config` option.

In the following command, `switches.lst` is a file that contains the switches to be queried.

```
$ ./patchmgr --roceswitches switches.lst --verify-config -log_dir /tmp
```

```
2019-11-20 14:12:27 -0800 :Working: Initiate config verify on RoCE
switches from . Expect up to 6 minutes for each switch
```

```
2019-11-20 14:12:30 -0800 1 of 15 :Verifying config on switch racklsw-
rocea0
```

```
2019-11-20 14:12:30 -0800: [INFO ] Dumping current running config
locally as file: /tmp/run.racklsw-rocea0.cfg
```

```
2019-11-20 14:12:33 -0800: [SUCCESS ] Backed up switch config
successfully
```

```
2019-11-20 14:12:33 -0800: [INFO ] Validating running config against
template [1/3]: /tmp/patch_switch_19.3.1.0.0.191018/
roce_switch_templates/roce_leaf_switch.cfg
```

```
2019-11-20 14:12:33 -0800: [INFO ] Validating running config against
template [2/3]: /tmp/patch_switch_19.3.1.0.0.191018/
roce_switch_templates/roce_leaf_switch_multi.cfg
```

```
2019-11-20 14:12:33 -0800: [INFO ] Config matches template: /tmp/
patch_switch_19.3.1.0.0.191018/roce_switch_templates/
roce_leaf_switch_multi.cfg
```

```
2019-11-20 14:12:33 -0800: [SUCCESS ] Config validation successful!
```

```
2019-11-20 14:12:33 -0800 2 of 15 :Verifying config on switch racklsw-
roceb0
```

```
...
```

7. Perform the physical cabling of the switches in the new rack R_{n+1} .

Caution:

Cabling within a live network must be done carefully in order to avoid potentially serious disruptions.

- a. Remove the eight existing inter-switch connections between each leaf switch in the new rack R_{n+1} (ports 4, 5, 6, 7 and 30, 31, 32, 33).

- b. Cable the leaf switches in the new rack according to the appropriate table in [Multi-Rack Cabling Tables](#).

For example, if you are adding a 5th rack and rack R_{n+1} is R5, then use "Table 5-14 Leaf Switch Connections for the Fifth Rack in a Five-Rack System".

8. Add the new rack to the switches in the existing racks (R1 to R_n).
 - a. For an existing rack (Rx), cable the lower leaf switch RxLL according to the appropriate table in [Multi-Rack Cabling Tables](#).
 - b. For the same rack, cable the upper leaf switch RxUL according to the appropriate table in [Multi-Rack Cabling Tables](#).
 - c. Repeat these steps for each existing rack, R1 to R_n .
9. Confirm each switch is available and connected.

For each switch in racks R1, R2, ..., R_n , R_{n+1} , confirm the output for the switch `show interface status` command shows `connected` and 100G. In the following example, the leaf switches are ports Eth1/4 to Eth1/7, and Eth1/30 to Eth1/33. The spine switches are ports Eth1/5 to Eth1/20.

When run from a spine switch, the output should be similar to the following:

```
rack1sw-roces0# show interface status
-----
-----
Port      Name          Status  Vlan  Duplex
Speed    Type
-----
-----
mgmt0     --           connected routed  full
1000     --
-----
-----
Port      Name          Status  Vlan  Duplex
Speed    Type
-----
-----
...
Eth1/5    RouterPort5   connected routed  full
100G     QSFP-100G-CR4
Eth1/6    RouterPort6   connected routed  full
100G     QSFP-100G-SR4
Eth1/7    RouterPort7   connected routed  full
100G     QSFP-100G-CR4
Eth1/8    RouterPort8   connected routed  full
100G     QSFP-100G-SR4
Eth1/9    RouterPort9   connected routed  full
100G     QSFP-100G-CR4
Eth1/10   RouterPort10  connected routed  full
100G     QSFP-100G-SR4
Eth1/11   RouterPort11  connected routed  full
100G     QSFP-100G-CR4
Eth1/12   RouterPort12  connected routed  full
100G     QSFP-100G-SR4
Eth1/13   RouterPort13  connected routed  full
```

```

100G   QSFP-100G-CR4
Eth1/14   RouterPort14   connected routed   full   100G
QSFP-100G-SR4
Eth1/15   RouterPort15   connected routed   full   100G
QSFP-100G-CR4
Eth1/16   RouterPort16   connected routed   full   100G
QSFP-100G-SR4
Eth1/17   RouterPort17   connected routed   full   100G
QSFP-100G-CR4
Eth1/18   RouterPort18   connected routed   full   100G
QSFP-100G-SR4
Eth1/19   RouterPort19   connected routed   full   100G
QSFP-100G-CR4
Eth1/20   RouterPort20   connected routed   full   100G
QSFP-100G-SR4
Eth1/21   RouterPort21   xcvrAbsen   routed   full   100G
--
...
  
```

When run from a leaf switch, the output should be similar to the following:

```

rack1sw-rocea0# show interface status
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
mgmt0         --            connected   routed    full    1000   --
-----
-----
Port          Name          Status      Vlan      Duplex  Speed  Type
-----
...
Eth1/4        RouterPort1   connected routed    full    100G
QSFP-100G-CR4
Eth1/5        RouterPort2   connected routed    full    100G
QSFP-100G-CR4
Eth1/6        RouterPort3   connected routed    full    100G
QSFP-100G-CR4
Eth1/7        RouterPort4   connected routed    full    100G
QSFP-100G-CR4
Eth1/8        celadm14     connected   3888     full    100G
QSFP-100G-CR4
...
Eth1/29       celadm01     connected   3888     full    100G
QSFP-100G-CR4
Eth1/30       RouterPort5   connected routed    full    100G
QSFP-100G-SR4
Eth1/31       RouterPort6   connected routed    full    100G
QSFP-100G-SR4
Eth1/32       RouterPort7   connected routed    full    100G
QSFP-100G-SR4
Eth1/33       RouterPort8   connected routed    full    100G
  
```

```
QSF-P-100G-SR4
...
```

10. Check the neighbor discovery for every switch in racks R1, R2, ..., Rn, Rn+1.

Log in to each switch and use the `show lldp neighbors` command. Make sure that all switches are visible and check the switch ports assignment (leaf switches: ports Eth1/4 - Eth1/7, Eth1/30 - Eth1/33; spine switches: ports Eth1/5 - Eth1/20) against the appropriate table in [Multi-Rack Cabling Tables](#).

Each spine switch should see all the leaf switches in each rack, but not the other spine switches. The output for a spine switch should be similar to the following:

 **Note:**

The interfaces in the rightmost output column (for example, Ethernet1/5) are different for each switch based on the applicable cabling tables.

```
rack1sw-roces0# show lldp neighbors | grep roce
rack1sw-roceb0 Eth1/5 120 BR Ethernet1/5
rack2sw-roceb0 Eth1/6 120 BR Ethernet1/5
rack1sw-roceb0 Eth1/7 120 BR Ethernet1/7
rack2sw-roceb0 Eth1/8 120 BR Ethernet1/7
rack1sw-roceb0 Eth1/9 120 BR Ethernet1/4
rack2sw-roceb0 Eth1/10 120 BR Ethernet1/4
rack3sw-roceb0 Eth1/11 120 BR Ethernet1/5
rack3sw-roceb0 Eth1/12 120 BR Ethernet1/7
rack1sw-rocea0 Eth1/13 120 BR Ethernet1/5
rack2sw-rocea0 Eth1/14 120 BR Ethernet1/5
rack1sw-rocea0 Eth1/15 120 BR Ethernet1/7
rack2sw-rocea0 Eth1/16 120 BR Ethernet1/7
rack3sw-rocea0 Eth1/17 120 BR Ethernet1/5
rack2sw-rocea0 Eth1/18 120 BR Ethernet1/4
rack3sw-rocea0 Eth1/19 120 BR Ethernet1/7
rack3sw-rocea0 Eth1/20 120 BR Ethernet1/4
```

Each leaf switch should see the spine switch in every rack, but not the other leaf switches. The output for a leaf switch should be similar to the following:

 **Note:**

The interfaces in the rightmost output column (for example, Ethernet1/13) are different for each switch based on the applicable cabling tables.

```
rack1sw-rocea0# show lldp neighbors | grep roce
rack3sw-roces0 Eth1/4 120 BR Ethernet1/13
rack1sw-roces0 Eth1/5 120 BR Ethernet1/13
rack3sw-roces0 Eth1/6 120 BR Ethernet1/15
```

```
rack1sw-roces0 Eth1/7 120 BR Ethernet1/15
rack2sw-roces0 Eth1/30 120 BR Ethernet1/17
rack2sw-roces0 Eth1/31 120 BR Ethernet1/13
rack3sw-roces0 Eth1/32 120 BR Ethernet1/17
rack2sw-roces0 Eth1/33 120 BR Ethernet1/15
```

11. Power on all the servers in the new rack, $Rn+1$.
12. For each rack, confirm the multi-rack cabling by running the `verify_roce_cables.py` script.

Refer to My Oracle Support Doc ID 2587717.1 for download and usage instructions.

Check the output of the `verify_roce_cables.py` script against the applicable tables in [Multi-Rack Cabling Tables for Oracle Exadata Rack X8M](#). Also, check that output in the `CABLE OK?` columns contains the `OK` status.

When running the script, two input files are used, one for nodes and one for switches. Each file should contain the servers or switches on separate lines. Use fully qualified domain names or IP addresses for each server and switch.

The following output is a partial example of the command results:

```
# ./verify_roce_cables.py -n nodes.rack1 -s switches.rack1
SWITCH PORT (EXPECTED PEER) LEAF-1 (rack1sw-rocea0) : CABLE OK? LEAF-2 (rack1sw-
roceb0) : CABLE OK?
-----
Eth1/4 (ISL peer switch) : rack1sw-roces0 Ethernet1/17 : OK rack1sw-roces0
Ethernet1/9 : OK
Eth1/5 (ISL peer switch) : rack1sw-roces0 Ethernet1/13 : OK rack1sw-roces0
Ethernet1/5 : OK
Eth1/6 (ISL peer switch) : rack1sw-roces0 Ethernet1/19 : OK rack1sw-roces0
Ethernet1/11: OK
Eth1/7 (ISL peer switch) : rack1sw-roces0 Ethernet1/15 : OK rack1sw-roces0
Ethernet1/7 : OK
Eth1/12 (celadm10) : rack1celadm10 port-1 : OK rack1celadm10
port-2 : OK
Eth1/13 (celadm09) : rack1celadm09 port-1 : OK rack1celadm09
port-2 : OK
Eth1/14 (celadm08) : rack1celadm08 port-1 : OK rack1celadm08
port-2 : OK
...
Eth1/15 (adm08) : rack1dbadm08 port-1 : OK rack1dbadm08
port-2 : OK
Eth1/16 (adm07) : rack1dbadm07 port-1 : OK rack1dbadm07
port-2 : OK
Eth1/17 (adm06) : rack1dbadm06 port-1 : OK rack1dbadm06
port-2 : OK
...
Eth1/30 (ISL peer switch) : rack2sw-roces0 Ethernet1/17 : OK rack2sw-roces0
Ethernet1/9 : OK
Eth1/31 (ISL peer switch) : rack2sw-roces0 Ethernet1/13 : OK rack2sw-roces0
Ethernet1/5 : OK
Eth1/32 (ISL peer switch) : rack2sw-roces0 Ethernet1/19 : OK rack2sw-roces0
Ethernet1/11: OK
```

```
Eth1/33 (ISL peer switch) : rack2sw-roces0 Ethernet1/15 : OK      rack2sw-
roces0 Ethernet1/7 : OK
```

13. Verify the RoCE Network Fabric operation by using the `infinicheck` command.

Use the following recommended command sequence. In each command, `hosts.lst` is the name of an input file that contains a comma-delimited list of database server host names or RoCE Network Fabric IP addresses, and `cells.lst` is the name of an input file that contains a list of RoCE Network Fabric IP addresses for the storage servers.

- Use `infinicheck` with the `-z` option to clear the files that were created during the last run of the `infinicheck` command. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -
c cells.lst -z
```

- Use `infinicheck` with the `-s` option to set up user equivalence for password-less SSH across the RoCE Network Fabric. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -
c cells.lst -s
```

- Finally, verify the RoCE Network Fabric operation by using `infinicheck` with the `-b` option, which is recommended on newly imaged machines where it is acceptable to suppress the `cellip.ora` and `cellinit.ora` configuration checks. For example:

```
# /opt/oracle.SupportTools/ibdiagtools/infinicheck -g hosts.lst -
c cells.lst -b
```

```
INFINICHECK
      [Network Connectivity, Configuration and
Performance]

      ##### FABRIC TYPE TESTS #####
System type identified: RoCE
Verifying User Equivalence of user=root from all DBs to all
CELLs.
      ##### RoCE CONFIGURATION TESTS #####
      Checking for presence of RoCE devices on all DBs and CELLS
[SUCCESS].... RoCE devices on all DBs and CELLS look good
      Checking for RoCE Policy Routing settings on all DBs and
CELLs
[SUCCESS].... RoCE Policy Routing settings look good
      Checking for RoCE DSCP ToS mapping on all DBs and CELLS
[SUCCESS].... RoCE DSCP ToS settings look good
      Checking for RoCE PFC settings and DSCP mapping on all DBs
and CELLS
[SUCCESS].... RoCE PFC and DSCP settings look good
      Checking for RoCE interface MTU settings. Expected value :
2300
[SUCCESS].... RoCE interface MTU settings look good
      Verifying switch advertised DSCP on all DBs and CELLS ports
( )
```

```
[SUCCESS].... Advertised DSCP settings from RoCE switch looks good
#### CONNECTIVITY TESTS ####
[COMPUTE NODES -> STORAGE CELLS]
(60 seconds approx.)
(Will walk through QoS values: 0-6) [SUCCESS].....Results OK
[SUCCESS]..... All can talk to all storage cells
[COMPUTE NODES -> COMPUTE NODES]
...

```

14. After cabling the racks together, proceed to [Configuring the New Hardware](#) to finish the configuration of the new rack.

Related Topics

- [Exadata Database Machine and Exadata Storage Server Supported Versions \(My Oracle Support Doc ID 888828.1\)](#)

2.3.3.3 Cabling Several InfiniBand Network Fabric Racks Together

Use this procedure to add another rack to an existing multi-rack system with InfiniBand Network Fabric.

This procedure is for systems with InfiniBand Network Fabric (X8, or earlier).

1. Set the priority of the current active Subnet Manager Master to 10 on a spine switch.
 - a. Log in to any InfiniBand switch on the active system.
 - b. Use the `getmaster` command to determine that the Subnet Manager Master is running on the spine switch.

The following example shows that the Subnet Manager Master is running on the spine switch `dm01sw-ib1`.

```
# getmaster
20100701 11:46:38 OpenSM Master on Switch : 0x0021283a8516a0a0 ports
36 Sun DCS 36
QDR switch dm01sw-ib1.example.com enhanced port 0 lid 1 lmc 0

```

If the Subnet Manager Master is not running on a spine switch, then perform the following steps:

- i. Use the `getmaster` command to identify the current location of the Subnet Manager Master.
 - ii. Log in as the `root` user on the leaf switch that is the Subnet Manager Master.
 - iii. Disable Subnet Manager on the switch. The Subnet Manager Master relocates to another switch.
 - iv. Use the `getmaster` command to identify the current location of the Subnet Manager Master. If a spine switch is not the Subnet Manager Master, then repeat steps [1.b.ii](#) and [1.b.iii](#) until a spine switch is the Subnet Manager Master.
 - v. Enable Subnet Manager on the leaf switches that were disabled during this procedure.
- c. Log in to the Subnet Manager Master spine switch.
 - d. Use the `disablesm` command to stop the Subnet Manager.

- e. Use the `setsmpriority 10` command to set the priority to 10.
 - f. Use the `enableesm` command to restart the Subnet Manager.
 - g. Repeat step 1.b to ensure that the Subnet Manager Master is running on the spine switch.
2. Ensure the new rack is near the existing rack.
The InfiniBand cables must be able to reach the servers in each rack.
 3. Completely shut down the new rack (R_{n+1}).
 4. Cable the leaf switch in the new rack according to the appropriate table in [Multi-Rack Cabling Tables for Oracle Exadata Rack Models with InfiniBand Network Fabric \(X2 to X8\)](#).

For example, if rack R_{n+1} is R4, then use [Table 6-9](#).

 **Caution:**

Cabling within a live network must be done carefully in order to avoid potentially serious disruptions.

The cabling table that you use for your new InfiniBand topology tells you how to connect ports on the leaf switches to ports on spine switches in order to connect the racks. Some of these ports on the spine switches might be already in use to support the existing InfiniBand topology. In these cases, connect only the cable on the leaf switch in the new rack and stop there for now. Make note of which cables you were not able to terminate.

Do not unplug any cables on the spine switch in the existing rack at this point. Step 5 describes how to re-cable the leaf switches on the existing racks (one leaf switch after the other - while the leaf switch being re-cabled will be powered off), which will free up these currently in-use ports. At that point, you can connect the other end of the cable from the leaf switch in the new rack to the spine switch in the existing rack as indicated in the table.

5. Complete the following procedure for each of the original racks:

In these steps, R_x represents a rack number from R1 to R_n .

- a. Power off leaf switch R_x IB2.
This causes all servers in the rack to fail over their InfiniBand traffic to R_x IB3.
- b. Cable leaf switch R_x IB2 according to [Multi-Rack Cabling Tables for Oracle Exadata Rack Models with InfiniBand Network Fabric \(X2 to X8\)](#).
- c. Power on leaf switch R_x IB2.
- d. Wait at least three minutes for R_x IB2 to become completely operational.
To check the switch, log in to the switch and run the `ibswitches` command. The output should show $n*3$ switches for IB1, IB2, and IB3 in racks R1, R2, ... R_n .
- e. Power off leaf switch R_x IB3.
This causes all servers in the rack to fail over their InfiniBand traffic to R_x IB2.

- f. Cable leaf switch Rx IB3 according to [Multi-Rack Cabling Tables for Oracle Exadata Rack Models with InfiniBand Network Fabric \(X2 to X8\)](#).
- g. Power on leaf switch Rx IB3.
- h. Wait at least three minutes for Rx IB3 to become completely operational.
To check the switch, log in to the switch and run the `ibswitches` command. The output should show $n*3$ switches for IB1, IB2, and IB3 in racks R1, R2, ... Rn.

All racks should now be rewired according to [Multi-Rack Cabling Tables for Oracle Exadata Rack Models with InfiniBand Network Fabric \(X2 to X8\)](#).

6. Power on all the InfiniBand switches in the new rack.
7. Wait three minutes for the switches to become completely operational.
To check the switch, log in to the switch and run the `ibswitches` command. The output should show $(n+1)*3$ switches for IB1, IB2, and IB3 in racks R1, R2, ... Rn+1.
8. Ensure that the Subnet Manager Master is running on R1 IB1 by running the `getmaster` command from any switch.
9. Power on all servers in the new rack (Rn+1).
10. Log in to spine switch R1 IB1, and lower its priority to 8.
 - a. Use the `disableesm` command to stop Subnet Manager.
 - b. Use the `setsmpriority 8` command to set the priority to 8.
 - c. Use the `enableesm` command to restart Subnet Manager.
11. Ensure that the Subnet Manager Master is running on one of the spine switches using the `getmaster` command from any switch.
12. Ensure that the Subnet Manager is running on every spine switch by entering the following command from any switch:

```
ibdiagnet -r
```

Each spine switch should show as running in the `Summary Fabric SM-state-priority` section of the output. If a spine switch is not running, then log in to the switch and enable the Subnet Manager using the `enableesm` command.

13. If there are now four or more racks, then log in to the leaf switches in each rack and disable Subnet Manager using the `disableesm` command.

3

Configuring the New Hardware

This section contains the following tasks needed to configure the new hardware:



Note:

The new and existing racks must be at the same patch level for Oracle Exadata Database Servers and Oracle Exadata Storage Servers, including the operating system. Refer to [Reviewing Release and Patch Levels](#) for additional information.

- [Changing the Interface Names](#)
- [Setting Up New Servers](#)
New servers need to be configured when extending Exadata Database Machine Elastic Configurations.
- [Setting up a New Rack](#)
A new rack is configured at the factory. However, it is necessary to set up the network and configuration files for use with the existing rack.
- [Setting User Equivalence](#)
User equivalence can be configured to include all servers once the servers are online.
- [Starting the Cluster](#)
- [Adding Grid Disks to Oracle ASM Disk Groups](#)
- [Adding Servers to a Cluster](#)
This procedure describes how to add servers to a cluster.
- [Configuring Cell Alerts for New Oracle Exadata Storage Servers](#)
Cell alerts need to be configured for the new Oracle Exadata Storage Servers.
- [Adding Oracle Database Software to the New Servers](#)
- [Adding Database Instance to the New Servers](#)
- [Returning the Rack to Service](#)

3.1 Changing the Interface Names

For systems with RoCE Network Fabric (X8M and later), `BONDETH0` is used for the bonded Ethernet client network.

For later systems with InfiniBand Network Fabric (X3 to X8), `BONDIB0` and `BONDETH0` are typically used for the bonded RDMA Network Fabric and the bonded Ethernet client network, respectively.

For Oracle Exadata Database Machine X2-2 and earlier systems (with X4170 and X4275 servers), `BOND0` and `BOND1` are the names for the bonded RDMA Network Fabric and bonded Ethernet client networks, respectively.

If you are adding new servers to an existing Exadata Database Machine, ensure the database servers use the same names for bonded configuration. You can either change the new database servers to match the existing server interface names, or change the existing server interface names and Oracle Cluster Registry (OCR) configuration to match the new servers.

Do the following after changing the interface names:

1. Edit the entries in `/etc/sysctl.conf` file on the database servers so that the entries for the RDMA Network Fabric match. The following is an example of the file entries before editing. One set of entries must be changed to match the other set.

Found in X2 node

```
net.ipv4.neigh.bondib0.locktime = 0
net.ipv4.conf.bondib0.arp_ignore = 1
net.ipv4.conf.bondib0.arp_accept = 1
net.ipv4.neigh.bondib0.base_reachable_time_ms = 10000
net.ipv4.neigh.bondib0.delay_first_probe_time = 1
```

Found in V2 node

```
net.ipv4.conf.bond0.arp_accept=1
net.ipv4.neigh.bond0.base_reachable_time_ms=10000
net.ipv4.neigh.bond0.delay_first_probe_time=1
```

2. Save the changes to the `sysctl.conf` file.
3. Use the `oifcfg` utility to change the OCR configuration, if the new names differ from what is currently in OCR. The interface names for Oracle Exadata Storage Servers do not have to be changed.
4. Continue configuring the new hardware, as follows:
 - If the hardware is new servers, then go to [Setting Up New Servers](#) to configure the servers.
 - If the hardware is a new rack, then go to [Setting up a New Rack](#) to configure the rack.



See Also:

Oracle Exadata Database Machine Maintenance Guide for information about changing the RDMA Network Fabric information

3.2 Setting Up New Servers

New servers need to be configured when extending Exadata Database Machine Elastic Configurations.

The new servers do not have any configuration information, and you cannot use Oracle Enterprise Manager Cloud Control to configure them. The servers are configured using the Oracle Exadata Deployment Assistant (OEDA) or manually.

- [Configuring Servers Using OEDA](#)
When adding servers to an Exadata Database Machine, you can use OEDA.
- [Configuring New Servers Manually](#)
When adding servers to an Exadata Database Machine, you can configure the servers manually instead of using OEDA.

3.2.1 Configuring Servers Using OEDA

When adding servers to an Exadata Database Machine, you can use OEDA.

**Note:**

In order to configure the servers with Oracle Exadata Deployment Assistant (OEDA), the new server information must be entered in OEDA, and configuration files generated.

1. Download the latest release of OEDA listed in My Oracle Support note 888828.1.
2. Enter the new server information in OEDA.

Do not include information for the existing rack.

**Note:**

- When extending an existing rack that has database servers earlier than Exadata Database Machine X4-2, be sure to deselect the active bonding option for the InfiniBand Network Fabric so the new database servers are configured with active-passive bonded interfaces.
- When extending an existing Exadata Database Machine X4-2 or later system with active-active bonding, select the active bonding option to configure the new database servers for active-active bonding.

3. Use OEDA to generate the configuration files.
4. Prepare the servers as follows, starting with the first database server of the new servers:
 - a. Configure the servers as described in [Preparing the Servers](#).

**Note:**

OEDA checks the performance level of Oracle Exadata Storage Servers so it is not necessary to check them using the CellCLI `CALIBRATE` command at this time.

- b. Create the cell disks and grid disks as described in [Configuring Cells, Cell Disks, and Grid Disks with CellCLI](#).
- c. Create the flash cache and flash log as described in [Creating Flash Cache and Flash Grid Disks](#).

 **Note:**

When creating the flash cache, enable write-back flash cache.

5. Ensure the RDMA Network Fabric and bonded client Ethernet interface names are the same on the new database servers as on the existing database servers.
6. When using the same, earlier style bonding names, such as `BOND0`, for the new database servers, then update the `/opt/oracle.cellos/cell.conf` file to reflect the correct bond names.

 **Note:**

If the existing servers use the latest bonded interface names, such as `BONDIB0`, then this step can be skipped.

7. Install OEDA on the first new database server.

 **See Also:**

My Oracle Support note 888828.1 for information about OEDA

8. Copy the configuration files to the first database server of the *new* servers in the `/opt/oracle.SupportTools/onecommand` directory.

This is the information completed in step 2.

9. Run OEDA up to, but not including, the `Create Grid Disks` step, and then run the `Configure Alerting` and `Setup ASR Alerting` configuration steps.

 **Note:**

- The `OEDA Validate Configuration File` step may display an error message about missing files, `pXX.zip`. This is expected behavior because the files are not used for this procedure. You can ignore the error message.
- When using capacity-on-demand, OEDA has the `Setup Capacity-on-Demand` step. This step uses the `setupResourceControl` command to set up the cores correctly.

10. Configure the storage servers, cell disks and grid disks as described in `Configuring Cells, Cell Disks, and Grid Disks with CellCLI`.

 **Note:**

Use the data collected from the existing system, as described in [Obtaining Current Configuration Information](#) to determine the grid disk names and sizes.

11. On systems with InfiniBand Network Fabric only, run `/opt/oracle.SupportTools/reclaimdisks.sh -free -reclaim` on each database server.

The command reclaims disk space reserved for the deployment type not selected, which is not required on RoCE-based Exadata database servers.

The command typically takes less than 5 minutes.

Do not skip this step. Skipping this step results in unused space that can no longer be reclaimed by `reclaimdisks.sh`.

12. Verify the time is the same on the new servers as on the existing servers.
This check is performed for the storage servers and database servers.
13. Ensure the NTP settings are the same on the new servers as on the existing servers.
This check is performed for the storage servers and database servers.
14. Configure HugePages on the new servers to match the existing servers.
15. Ensure the values in the `/etc/security/limits.conf` file for the new database servers match the existing database servers.
16. Go to [Setting User Equivalence](#) to continue the hardware configuration.

Related Topics

- [Preparing the Network Configuration](#)
When adding additional servers to your rack, you will need IP address and the current network configuration settings.
- [Exadata Database Machine and Exadata Storage Server Supported Versions \(My Oracle Support Doc ID 888828.1\)](#)

3.2.2 Configuring New Servers Manually

When adding servers to an Exadata Database Machine, you can configure the servers manually instead of using OEDA.

Using Oracle Exadata Deployment Assistant (OEDA) is the recommended method.

1. Prepare the servers using the procedure described in [Preparing the Servers](#).
2. Ensure the RDMA Network Fabric and bonded client Ethernet interface names are the same on the new database servers as on the existing database servers.
3. Configure the storage servers, cell disks and grid disks as described in [Configuring Cells, Cell Disks, and Grid Disks with CellCLI](#).
4. Configure the database servers as described in [Setting Up Configuration Files for a Database Server Host](#).
5. On systems with InfiniBand Network Fabric only, run `/opt/oracle.SupportTools/reclaimdisks.sh -free -reclaim` on each database server.

The command reclaims disk space reserved for the deployment type not selected, which is not required on RoCE-based Exadata database servers.

The command typically takes less than 5 minutes.

Do not skip this step. Skipping this step results in unused space that can no longer be reclaimed by `reclaimdisks.sh`.

6. Verify the time is the same on the new servers as on the existing servers.
This check is performed for the storage servers and database servers.
7. Ensure the NTP settings are the same on the new servers as on the existing servers.
This check is performed for the storage servers and database servers.
8. Configure HugePages on the new servers to match the existing servers.
9. Go to [Setting User Equivalence](#) to continue the hardware configuration.

Related Topics

- [Preparing the Network Configuration](#)
When adding additional servers to your rack, you will need IP address and the current network configuration settings.
- [Exadata Database Machine and Exadata Storage Server Supported Versions \(My Oracle Support Doc ID 888828.1\)](#)

3.3 Setting up a New Rack

A new rack is configured at the factory. However, it is necessary to set up the network and configuration files for use with the existing rack.

1. Configure the new rack and servers as described in *Configuring Oracle Exadata Database Machine* in the *Oracle Exadata Database Machine Installation and Configuration Guide*.
Complete the appropriate tasks to configure the rack and its components, but do not complete the task for installing the Exadata configuration information and software on the servers. This task will be completed later in this procedure.
2. Verify the time is the same on the new servers as on the existing servers. This check is performed for storage servers and database servers.
3. Ensure the NTP settings are the same on the new servers as on the existing servers. This check is performed for storage servers and database servers.
4. Configure HugePages on the new servers to match the existing servers.
5. Ensure the RDMA Network Fabric and bonded client Ethernet interface names on the new database servers match the existing database servers.
6. Configure the rack as described in *Loading the Configuration Information and Installing the Software* in *Oracle Exadata Database Machine Installation and Configuration Guide*. You can use either the Oracle Exadata Deployment Assistant (OEDA) or Oracle Enterprise Manager Cloud Control to configure the rack.

 **Note:**

- Only run OEDA up to the `Create Grid Disks` step, then configure storage servers as described in *Configuring Cells, Cell Disks, and Grid Disks with CellCLI* in *Oracle Exadata System Software User's Guide*.
- When adding servers with 3 TB High Capacity (HC) disks to existing servers with 2TB disks, it is recommended to follow the procedure in My Oracle Support Doc ID 1476336.1 to properly define the grid disks and disk groups. At this point of setting up the rack, it is only necessary to define the grid disks. The disk groups are created after the cluster has been extended on to the new nodes.
- If the existing storage servers are Extreme Flash (EF) and you are adding High Capacity (HC) storage servers, or if the existing storage servers are HC and you are adding EF storage servers, then you must place the new disks in new disk groups. You cannot mix EF and HC disks within the same disk group.

7. Go to [Setting User Equivalence](#) to continue the hardware configuration.

Related Topics

- [How to Add Exadata Storage Servers Using 3 TB \(or larger\) Disks to an Existing Database Machine \(My Oracle Support Doc ID 1476336.1\)](#)

3.4 Setting User Equivalence

User equivalence can be configured to include all servers once the servers are online.

This procedure must be done before running the post-cabling utilities.

1. Log in to each new server manually using SSH to verify that each server can accept log ins and that the passwords are correct.
2. Modify the `dbs_group` and `cell_group` files on all servers to include all servers.

- a. Create the new directories on the first existing database server.

```
# mkdir /root/new_group_files
# mkdir /root/old_group_files
# mkdir /root/group_files
```

- b. Copy the group files for the new servers to the `/root/new_group_files` directory.
- c. Copy the group files for the existing servers to the `/root/old_group_files` directory.
- d. Copy the group files for the existing servers to the `/root/group_files` directory.
- e. Update the group files to include the existing and new servers.

```
cat /root/new_group_files/dbs_group >> /root/group_files/dbs_group
cat /root/new_group_files/cell_group >> /root/group_files/cell_group
cat /root/new_group_files/all_group >> /root/group_files/all_group
cat /root/new_group_files/dbs_priv_group >> /root/group_files/
```

```
db_s_priv_group
cat /root/new_group_files/cell_priv_group >> /root/group_files/
cell_priv_group
cat /root/new_group_files/all_priv_group >> /root/group_files/
all_priv_group
```

- f. Make the updated group files the default group files. The updated group files contain the existing and new servers.

```
cp /root/group_files/* /root
cp /root/group_files/* /opt/oracle.SupportTools/onecommand
```

- g. Put a copy of the updated group files in the `root` user, `oracle` user, and Oracle Grid Infrastructure user home directories, and ensure that the files are owned by the respective users.
3. Modify the `/etc/hosts` file on the existing and new database server to include the existing RDMA Network Fabric IP addresses for the database servers and storage servers. The existing and new `all_priv_group` files can be used for this step.

 **Note:**

Do not copy the `/etc/hosts` file from one server to the other servers. Edit the file on each server.

4. Run the `setssh-Linux.sh` script as the `root` user on one of the existing database servers to configure user equivalence for all servers using the following command. Oracle recommends using the first database server.

```
# /opt/oracle.SupportTools/onecommand/setssh-Linux.sh -s -c N -h \  
/path_to_file/all_group -n N
```

In the preceding command, *path_to_file* is the directory path for the `all_group` file containing the names for the existing and new servers.

 **Note:**

For Oracle Exadata Database Machine X2-2 (with X4170 and X4275 servers) systems, use the `setssh.sh` command to configure user equivalence.

The command line options for the `setssh.sh` command differ from the `setssh-Linux.sh` command. Run `setssh.sh` without parameters to see the proper syntax.

5. Add the known hosts using RDMA Network Fabric. This step requires that all database servers are accessible by way of their InfiniBand interfaces.

```
# /opt/oracle.SupportTools/onecommand/setssh-Linux.sh -s -c N -h \
/path_to_file/all_priv_group -n N -p password
```

6. Verify equivalence is configured.

```
# dcli -g all_group -l root date
# dcli -g all_priv_group -l root date
```

7. Run the `setssh-Linux.sh` script as the `oracle` user on one of the existing database servers to configure user equivalence for all servers using the following command. Oracle recommends using the first database server. If there are separate owners for the Oracle Grid Infrastructure software, then run a similar command for each owner.

```
$ /opt/oracle.SupportTools/onecommand/setssh-Linux.sh -s -c N -h \
/path_to_file/dbs_group -n N
```

In the preceding command, *path_to_file* is the directory path for the `dbs_group` file. The file contains the names for the existing and new servers.

Note:

- For Oracle Exadata Database Machine X2-2 (with X4170 and X4275 servers) systems, use the `setssh.sh` command to configure user equivalence.
- It may be necessary to temporarily change the permissions on the `setssh-Linux.sh` file to 755 for this step. Change the permissions back to the original settings after completing this step.

8. Add the known hosts using RDMA Network Fabric. This step requires that all database servers are accessible by way of their InfiniBand interfaces.

```
$ /opt/oracle.SupportTools/onecommand/setssh-Linux.sh -s -c N -h \
/root/group_files/dbs_priv_group -n N
```

9. Verify equivalence is configured.

```
$ dcli -g dbs_group -l oracle date
$ dcli -g dbs_priv_group -l oracle date
```

If there is a separate Oracle Grid Infrastructure user, then also run the preceding commands for that user, substituting the `grid` user name for the `oracle` user.

3.5 Starting the Cluster

The following procedure describes how to start the cluster if it was stopped earlier for cabling an additional rack.

 **Note:**

- Oracle recommends you start one server, and let it come up fully before starting Oracle Clusterware on the rest of the servers.
- It is not necessary to stop a cluster when extending Oracle Exadata Database Machine Half Rack to a Full Rack, or a Quarter Rack to a Half Rack or Full Rack.

1. Log in as the `root` user on the original cluster.
2. Start one server of the cluster.

```
# Grid_home/grid/bin/crsctl start cluster
```

3. Check the status of the server.

```
Grid_home/grid/bin/crsctl stat res -t
```

Run the preceding command until it shows that the first server has started.

4. Start the other servers in the cluster.

```
# Grid_home/grid/bin/crsctl start cluster -all
```

5. Check the status of the servers.

```
Grid_home/grid/bin/crsctl stat res -t
```

It may take several minutes for all servers to start and join the cluster.

3.6 Adding Grid Disks to Oracle ASM Disk Groups

Grid disks can be added to Oracle ASM disk groups before or after the new servers are added to the cluster. The advantage of adding the grid disks before adding the new servers is that the rebalance operation can start earlier. The advantage of adding the grid disks after adding the new servers is that the rebalance operation can be done on the new servers so less load is placed on the existing servers.

The following procedure describes how to add grid disk to existing Oracle ASM disk groups.

 **Note:**

- It is assumed in the following examples that the newly-installed storage servers have the same grid disk configuration as the existing storage servers, and that the additional grid disks will be added to existing disk groups.

The information gathered about the current configuration should be used when setting up the grid disks.

- If the existing storage servers have High Performance (HP) disks and you are adding storage servers with High Capacity (HC) disks or the existing storage servers have HC disks and you are adding storage servers HP disks, then you must place the new disks in new disk groups. It is not permitted to mix HP and HC disks within the same disk group.

1. Ensure the new storage servers are running the same version of software as storage servers already in use. Run the following command on the first database server:

```
dcli -g dbs_group -l root "imageinfo -ver"
```

 **Note:**

If the Oracle Exadata System Software on the storage servers does not match, then upgrade or patch the software to be at the same level. This could be patching the existing servers or new servers. Refer to [Reviewing Release and Patch Levels](#) for additional information.

2. Modify the `/etc/oracle/cell/network-config/cellip.ora` file on all database servers to have a complete list of all storage servers. The `cellip.ora` file should be identical on all database servers.

When adding Oracle Exadata Storage Server X4-2L servers, the `cellip.ora` file contains two IP addresses listed for each cell. Copy each line completely to include the two IP addresses, and merge the addresses in the `cellip.ora` file of the existing cluster.

- a. From any database server, make a backup copy of the `cellip.ora` file.

```
cp /etc/oracle/cell/network-config  
cp cellip.ora cellip.ora.orig  
cp cellip.ora cellip.ora-bak
```

- b. Edit the `cellip.ora-bak` file and add the IP addresses for the new storage servers.
- c. Copy the edited file to the `cellip.ora` file on all database nodes using `dcli`. Use a file named `dbnodes` that contains the names of every database server in the cluster, with each database name on a separate line. Run the following command from the directory that contains the `cellip.ora-bak` file.

```
/usr/local/bin/dcli -g dbnodes -l root -f cellip.ora-bak -d  
/etc/oracle/cell/network-config/cellip.ora
```

The following is an example of the `cellip.ora` file after expanding Oracle Exadata Database Machine X3-2 Half Rack to Full Rack using Oracle Exadata Storage Server X4-2L servers:

```
cell="192.168.10.9"  
cell="192.168.10.10"  
cell="192.168.10.11"  
cell="192.168.10.12"  
cell="192.168.10.13"  
cell="192.168.10.14"  
cell="192.168.10.15"  
cell="192.168.10.17;192.168.10.18"  
cell="192.168.10.19;192.168.10.20"  
cell="192.168.10.21;192.168.10.22"  
cell="192.168.10.23;192.168.10.24"  
cell="192.168.10.25;192.168.10.26"  
cell="192.168.10.27;192.168.10.28"  
cell="192.168.10.29;192.168.10.30"
```

In the preceding example, lines 1 through 7 are for the original servers, and lines 8 through 14 are for the new servers. Oracle Exadata Storage Server X4-2L servers have two IP addresses each.

3. Ensure the updated `cellip.ora` file is on all database servers. The updated file must include a complete list of all storage servers.
4. Verify accessibility of all grid disks from one of the original database servers. The following command can be run as the `root` user or the `oracle` user.

```
$ Grid_home/grid/bin/kfod disks=all dscvgroup=true
```

The output from the command shows grid disks from the original and new storage servers.

5. Add the grid disks from the new storage servers to the existing disk groups using commands similar to the following. You cannot have both high performance disks and high capacity disks in the same disk group.

```
$.oraenv  
ORACLE_SID = [oracle] ? +ASM1  
The Oracle base for ORACLE_HOME=/u01/app/11.2.0/grid is /u01/app/  
oracle  
  
$ sqlplus / as sysasm  
SQL> ALTER DISKGROUP data ADD DISK  
2> 'o/*/DATA*dm02*'  
3> rebalance power 11;
```

In the preceding commands, a Full Rack was added to an existing Oracle Exadata Rack. The prefix for the new rack is `dm02`, and the grid disk prefix is `DATA`.

The following is an example in which an Oracle Exadata Database Machine Half Rack was upgraded to a Full Rack. The cell host names in the original system

were named dm01cel101 through dm01cel107. The new cell host names are dm01cel108 through dm01cel114.

```
$ .oraenv
ORACLE_SID = [oracle] ? +ASM1
The Oracle base for ORACLE_HOME=/u01/app/11.2.0/grid is /u01/app/oracle

$ SQLPLUS / AS sysasm
SQL> ALTER DISKGROUP data ADD DISK
 2> 'o/*/DATA*dm01cel108*',
 3> 'o/*/DATA*dm01cel109*',
 4> 'o/*/DATA*dm01cel110*',
 5> 'o/*/DATA*dm01cel111*',
 6> 'o/*/DATA*dm01cel112*',
 7> 'o/*/DATA*dm01cel113*',
 8> 'o/*/DATA*dm01cel114*'
 9> rebalance power 11;
```

 **Note:**

- If your system is running Oracle Database 11g release 2 (11.2.0.1), then Oracle recommends a power limit of 11 so that the rebalance completes as quickly as possible. If your system is running Oracle Database 11g release 2 (11.2.0.2), then Oracle recommends a power limit of 32. The power limit does have an impact on any applications that are running during the rebalance.
 - Ensure the `ALTER DISKGROUP` commands are run from different Oracle ASM instances. That way, the rebalance operation for multiple disk groups can run in parallel.
 - Add disks to all disk groups including `SYSTEMDG` or `DBFS_DG`.
 - When adding servers with 3 TB High Capacity (HC) disks to existing servers with 2 TB disks, it is recommended to follow the procedure in My Oracle Support note 1476336.1 to properly define the grid disks and disk groups. At this point of setting up the rack, the new grid disks should be defined, but need to be placed into disk groups. Refer to the steps in My Oracle Support note 1476336.1.
 - If the existing storage servers have High Performance (HP) disks and you are adding storage servers with High Capacity (HC) disks, or the existing storage servers have HC disks and you are adding storage servers with HP disks, then you must place the new disks in new disk groups. It is not permitted to mix HP and HC disks within the same disk group.
6. Monitor the status of the rebalance operation using a query similar to the following from any Oracle ASM instance:

```
SQL> SELECT * FROM GV$ASM_OPERATION WHERE STATE = 'RUN';
```

The remaining tasks can be done while the rebalance is in progress.

 **See Also:**

- [Obtaining Current Configuration Information](#) for information about the existing grid disks.
- [Setting Up New Servers](#) for information about configuring the grid disks.
- *Oracle Automatic Storage Management Administrator's Guide* for information about the `ASM_POWER_LIMIT` parameter.

3.7 Adding Servers to a Cluster

This procedure describes how to add servers to a cluster.

For adding nodes to an Oracle VM cluster, refer to Expanding an Oracle VM RAC Cluster on Exadata in *Oracle Exadata Database Machine Maintenance Guide*.

 **Caution:**

If Oracle Clusterware manages additional services that are not yet installed on the new nodes, such as Oracle GoldenGate, then note the following:

- It may be necessary to stop those services on the existing node before running the `addNode.sh` script.
- It is necessary to create any users and groups on the new database servers that run these additional services.
- It may be necessary to disable those services from auto-start so that Oracle Clusterware does not try to start the services on the new nodes.

 **Note:**

To prevent problems with transferring files between existing and new nodes, you need to set up SSH equivalence. See Step 4 in Expanding an Oracle VM Oracle RAC Cluster on Exadata in for details.

1. Ensure the `/etc/oracle/cell/network-config/*.ora` files are correct and consistent on all database servers. The `cellip.ora` file all database server should include the older and newer database servers and storage servers.
2. Ensure the `ORACLE_BASE` and `diag` destination directories have been created on the Oracle Grid Infrastructure destination home.

The following is an example for Oracle Grid Infrastructure 11g:

```
# dcli -g /root/new_group_files/dbs_group -l root mkdir -p \
  /u01/app/11.2.0/grid /u01/app/oraInventory /u01/app/grid/diag

# dcli -g /root/new_group_files/dbs_group -l root chown -R
```

```
grid:oinstall \
/u01/app/11.2.0 /u01/app/oraInventory /u01/app/grid

# dcli -g /root/new_group_files/dbs_group -l root chmod -R 770 \
/u01/app/oraInventory

# dcli -g /root/new_group_files/dbs_group -l root chmod -R 755 \
/u01/app/11.2.0 /u01/app/11.2.0/grid
```

The following is an example for Oracle Grid Infrastructure 12c:

```
# cd /
# rm -rf /u01/app/*
# mkdir -p /u01/app/12.1.0.2/grid
# mkdir -p /u01/app/oracle/product/12.1.0.2/dbhome_1
# chown -R oracle:oinstall /u01
```

3. Ensure the `inventory` directory and Grid home directories have been created and have the proper permissions. The directories should be owned by the Grid user and the `OINSTALL` group. The `inventory` directory should have 770 permission, and the Oracle Grid Infrastructure home directories should have 755.

If you are running Oracle Grid Infrastructure 12c or later:

- Make sure `oraInventory` does not exist inside `/u01/app`.
 - Make sure `/etc/oraInst.loc` does not exist.
4. Create users and groups on the new nodes with the same user identifiers and group identifiers as on the existing nodes.

 **Note:**

If Oracle Exadata Deployment Assistant (OEDA) was used earlier, then these users and groups should have been created. Check that they do exist, and have the correct UID and GID values.

5. Log in as the Grid user on an existing host.
6. Verify the Oracle Cluster Registry (OCR) backup exists.

```
ocrconfig -showbackup
```

7. Verify that the additional database servers are ready to be added to the cluster using commands similar to following:

```
$ cluvfy stage -post hwos -n \
dm02db01, dm02db02, dm02db03, dm02db04, dm02db05, dm02db06, dm02db07, dm02db08
\
-verbose

$ cluvfy comp peer -refnode dm01db01 -n \
dm02db01, dm02db02, dm02db03, dm02db04, dm02db05, dm02db06, dm02db07, dm02db08
\
-orainv oinstall -osdba dba | grep -B 3 -A 2 mismatched
```

```
$ cluvfy stage -pre nodeadd -n \  
dm02db01, dm02db02, dm02db03, dm02db04, dm02db05, dm02db06, dm02db07, dm02d  
b08 \  
-verbose -fixup -fixupdir /home/grid_owner_name/fixup.d
```

In the preceding commands, *grid_owner_name* is the name of the Oracle Grid Infrastructure software owner, *dm02db01* through *db02db08* are the new database servers, and *refnode* is an existing database server.

 **Note:**

- The second and third commands do not display output if the commands complete correctly.
- An error about a voting disk, similar to the following, may be displayed:

```
ERROR:
PRVF-5449 : Check of Voting Disk location "o/192.168.73.102/
\
DATA_CD_00_dm01cel07(o/192.168.73.102/DATA_CD_00_dm01cel07)"
\
failed on the following nodes:
Check failed on nodes:
    dm01db01
    dm01db01:No such file or directory
...
PRVF-5431 : Oracle Cluster Voting Disk configuration check
```

If such an error occurs:

- If you are running Oracle Grid Infrastructure 11g, set the environment variable as follows:

```
$ export IGNORE_PREADDDNODE_CHECKS=Y
```

Setting the environment variable does not prevent the error when running the `cluvfy` command, but it does allow the `addNode.sh` script to complete successfully.

- If you are running Oracle Grid Infrastructure 12c or later, use the following `addnode` parameters: `-ignoreSysPrereqs -ignorePrereq`

In Oracle Grid Infrastructure 12c and later, `addnode` does not use the `IGNORE_PREADDDNODE_CHECKS` environment variable.

- If a database server was installed with a certain image and subsequently patched to a later image, then some operating system libraries may be older than the version expected by the `cluvfy` command. This causes the `cluvfy` command and possibly the `addNode.sh` script to fail.

It is permissible to have an earlier version as long as the difference in versions is minor. For example, `glibc-common-2.5-81.e15_8.2` versus `glibc-common-2.5-49`. The versions are different, but both are at version 2.5, so the difference is minor, and it is permissible for them to differ.

Set the environment variable `IGNORE_PREADDDNODE_CHECKS=Y` before running the `addNode.sh` script or use the `addnode` parameters `-ignoreSysPrereqs -ignorePrereq` with the `addNode.sh` script to workaround this problem.

8. Ensure that all directories inside the Oracle Grid Infrastructure home on the existing server have their executable bits set. Run the following commands as the root user.

```
find /u01/app/11.2.0/grid -type d -user root ! -perm /u+x ! \
\
-perm /g+x ! -perm o+x

find /u01/app/11.2.0/grid -type d -user grid_owner_name ! -
perm /u+x ! \
-perm /g+x ! -perm o+x
```

In the preceding commands, *grid_owner_name* is the name of the Oracle Grid Infrastructure software owner, and */u01/app/11.2.0/grid* is the Oracle Grid Infrastructure home directory.

If any directories are listed, then ensure the group and others permissions are *+x*. The *Grid_home/network/admin/samples*, *\$GI_HOME/crf/admin/run/crfmond*, and *Grid_home/crf/admin/run/crflogd* directories may need the *+x* permissions set.

If you are running Oracle Grid Infrastructure 12c or later, run commands similar to the following:

```
# chmod -R u+x /u01/app/12.1.0.2/grid/gpnp/gpnp_bcp*

# chmod -R o+rx /u01/app/12.1.0.2/grid/gpnp/gpnp_bcp*

# chmod o+r /u01/app/12.1.0.2/grid/bin/oradaemonagent /u01/app/
12.1.0.2/grid/srvm/admin/logging.properties

# chmod a+r /u01/app/oracle/product/12.1.0.2/dbhome_1/bin/*0

# chmod a+r /u01/app/oracle/product/12.1.0.2/dbhome_1/bin/*0

# chown -f gi_owner_name:dba /u01/app/12.1.0.2/grid/OPatch/ocm/bin/
emocmrsp
```

The *Grid_home/network/admin/samples* directory needs the *+x* permission:

```
chmod -R a+x /u01/app/12.1.0.2/grid/network/admin/samples
```

9. Run the following command. It is assumed that the Oracle Grid Infrastructure home is owned by the Grid user.

```
$ dcli -g old_db_nodes -l root chown -f grid_owner_name:dba \
/u01/app/11.2.0/grid/OPatch/ocm/bin/emocmrsp
```

10. This step is needed only if you are running Oracle Grid Infrastructure 11g. In Oracle Grid Infrastructure 12c, no response file is needed because the values are specified on the command line.

Create a response file, `add-cluster-nodes.rsp`, as the Grid user to add the new servers similar to the following:

```
RESPONSEFILE_VERSION=2.2.1.0.0

CLUSTER_NEW_NODES={dm02db01,dm02db02, \
dm02db03,dm02db04,dm02db05,dm02db06,dm02db07,dm02db08}

CLUSTER_NEW_VIRTUAL_HOSTNAMES={dm0201-vip,dm0202-vip,dm0203-vip,dm0204-
vip, \
dm0205-vip,dm0206-vip,dm0207-vip,dm0208-vip}
```

In the preceding file, the host names `dm02db01` through `dm02db08` are the new nodes being added to the cluster.

 **Note:**

The lines listing the server names should appear on one continuous line. They are wrapped in the documentation due to page limitations.

11. Ensure most of the files in the `Grid_home/rdbms/audit` and `Grid_home/log/diag/*` directories have been moved or deleted before extending a cluster.
12. Refer to My Oracle Support note 744213.1 if the installer runs out of memory. The note describes how to edit the `Grid_home/oui/ora-param.ini` file, and change the `JRE_MEMORY_OPTIONS` parameter to `-Xms512m-Xmx2048m`.
13. Add the new servers by running the `addNode.sh` script from an existing server as the Grid user.

- If you are running Oracle Grid Infrastructure 11g:

```
$ cd Grid_home/oui/bin
$ ./addNode.sh -silent -responseFile /path/to/add-cluster-nodes.rsp
```

- If you are running Oracle Grid Infrastructure 12c or later, run the `addnode.sh` command with the `CLUSTER_NEW_NODES` and `CLUSTER_NEW_VIRTUAL_HOSTNAMES` parameters. The syntax is:

```
$ ./addnode.sh -silent
"CLUSTER_NEW_NODES={comma_delimited_new_nodes}"
"CLUSTER_NEW_VIRTUAL_HOSTNAMES={comma_delimited_new_node_vips}"
```

For example:

```
$ cd Grid_home/addnode/

$ ./addnode.sh -silent
"CLUSTER_NEW_NODES={dm02db01,dm02db02,dm02db03,dm02db04,dm02db05,
dm02db06,dm02db07,dm02db08}" "CLUSTER_NEW_VIRTUAL_HOSTNAMES={dm02db01-
vip,dm02db02-vip,
dm02db03-vip,dm02db04-vip,dm02db05-vip,dm02db06-vip,dm02db07-
```

```
vip, dm02db08-vip}"  
-ignoreSysPrereqs -ignorePrereq
```

14. Verify the grid disks are visible from each of the new database servers.

```
$ Grid_home/grid/bin/kfod disks=all dscvgroup=true
```

15. Run the `orainstRoot.sh` script as the `root` user when prompted using the `dcli` utility.

```
$ dcli -g new_db_nodes -l root \  
/u01/app/oraInventory/orainstRoot.sh
```

16. Disable HAIP on the new servers.

Before running the `root.sh` script, on each new server, set the `HAIP_UNSUPPORTED` environment variable to `TRUE`.

```
$ export HAIP_UNSUPPORTED=true
```

17. Run the `Grid_home/root.sh` script on each server sequentially. This simplifies the process, and ensures that any issues can be clearly identified and addressed.

 **Note:**

The node identifier is set in order of the nodes where the `root.sh` script is run. Typically, the script is run from the lowest numbered node name to the highest.

18. Check the log file from the `root.sh` script and verify there are no problems on the server before proceeding to the next server. If there are problems, then resolve them before continuing.
19. Check the status of the cluster after adding the servers.

```
$ cluvfy stage -post nodeadd -n \  
  
dm02db01, dm02db02, dm02db03, dm02db04, dm02db05, dm02db06, dm02db07, dm02d  
b08 \  
-verbose
```

20. Check that all servers have been added and have basic services running.

```
crsctl stat res -t
```

 **Note:**

It may be necessary to mount disk groups on the new servers. The following commands must be run as the `oracle` user.

```
$ srvctl start diskgroup -g data
$ srvctl start diskgroup -g reco
```

21. If you are running Oracle Grid Infrastructure releases 11.2.0.2 and later, then perform the following steps:

- a. Manually add the `CLUSTER_INTERCONNECTS` parameter to the SPFILE for each Oracle ASM instance.

```
ALTER SYSTEM SET cluster_interconnects = '192.168.10.x' \
sid='+ASMx' scope=spfile
```

- b. Restart the cluster on each new server.
- c. Verify the parameters were set correctly.

3.8 Configuring Cell Alerts for New Oracle Exadata Storage Servers

Cell alerts need to be configured for the new Oracle Exadata Storage Servers.

The configuration depends on the type of installation.

- When extending Oracle Exadata Database Machine Quarter Rack to Half Rack, or Half Rack to Full Rack:

Manually configure cell alerts on the new storage servers. Use the settings on the original storage servers as a guide. To view the settings on the original storage servers, use a command similar to the following:

```
dcli -g new_cells_nodes -l celladmin cellcli -e list cell detail
```

To view the alerts on the new storage servers, use a command similar to the following:

```
dcli -g new_cell_nodes -l root "cellcli -e ALTER CELL \
smtpServer='\mailserver.example.com\' \
smtpPort=25, \
smtpUseSSL=false,smtpFrom='\DBM dm01\' \
smtpFromAddr='\storecell@example.com\' \
smtpToAddr='\dbm-admins@example.com\' \
notificationMethod='\mail,snmp\' \
notificationPolicy='\critical,warning,clear\' \
snmpSubscriber='\(host='\snmpserver.example.com, port=162\')\'"
```

 **Note:**

The backslash character (\) is used as an escape character for the dcli utility, and as a line continuation character in the preceding command.

- When cabling racks:
Use Oracle Exadata Deployment Assistant (OEDA) to set up e-mail alerts for storage servers as the `root` user from the original rack to the new rack. The utility includes the `SetupCellEmailAlerts` step to configure alerts.

3.9 Adding Oracle Database Software to the New Servers

It is necessary to add the Oracle Database software directory `ORACLE_HOME` to the database servers after the cluster modifications are complete, and all the database servers are in the cluster.

1. Check the `Oracle_home/bin` directory for files ending in zero (0), such as `nmb0`, that are owned by the `root` user and do not have `oinstall` or world read privileges. Use the following command to modify the file privileges:

```
# chmod a+r $ORACLE_HOME/bin/*0
```

If you are running Oracle Database release 12c or later, you also have to change permissions for files ending in uppercase O, in addition to files ending in zero.

```
# chmod a+r $ORACLE_HOME/bin/*O
```

2. This step is required for Oracle Database 11g only. If you are running Oracle Database 12c, you can skip this step because the directory has already been created.

Create the `ORACLE_BASE` directory for the database owner, if it is different from the Oracle Grid Infrastructure software owner (Grid user) using the following commands:

```
# dcli -g root/new_group_files/dbs_group -l root mkdir -p /u01/app/oracle
# dcli -g root/new_group_files/dbs_group -l root chown
oracle:oinstall \
/u01/app/oracle
```

3. Run the following command to set ownership of the `emocmrsp` file in the Oracle Database `$ORACLE_HOME` directory:

```
# dcli -g old_db_nodes -l root chown -f oracle:dba \
/u01/app/11.2.0/grid/OPatch/ocm/bin/emocmrsp
```

4. This step is required for Oracle Database 11g only. If you are running Oracle Database 12c, then you can skip this step because the values are entered on the command line.

Create a response file, `add-db-nodes.rsp`, as the `oracle` owner to add the new servers similar to the following:

```
RESPONSEFILE_VERSION=2.2.1.0.0

CLUSTER_NEW_NODES={dm02db01, dm02db02, dm02db03, dm02db04, dm02db05, \
dm02db06, dm02db07, dm02db08}
```

 **Note:**

The lines listing the server names should appear on one continuous line. The are wrapped in the document due to page limitations.

5. Add the Oracle Database `ORACLE_HOME` directory to the new servers by running the `addNode.sh` script from an existing server as the database owner user.

- If you are running Oracle Grid Infrastructure 11g:

```
$ cd $ORACLE_HOME/oui/bin
$ ./addNode.sh -silent -responseFile /path/to/add-db-nodes.rsp
```

- If you are running Oracle Grid Infrastructure 12c, then you specify the nodes on the command line. The syntax is:

```
./addnode.sh -silent "CLUSTER_NEW_NODES={comma_delimited_new_nodes}"
```

For example:

```
$ cd $Grid_home/addnode
$ ./addnode.sh -silent
"CLUSTER_NEW_NODES={dm02db01, dm02db02, dm02db03, dm02db04, dm02db05,
dm02db06, dm02db07, dm02db08}" -ignoreSysPrereqs -ignorePrereq
```

6. Ensure the `$ORACLE_HOME/oui/oraparam.ini` file has the memory settings that match the parameters set in the Oracle Grid Infrastructure home.
7. Run the `root.sh` script on each server when prompted as the `root` user using the `dcli` utility.

```
$ dcli -g new_db_nodes -l root $ORACLE_HOME/root.sh
```

In the preceding command, `new_db_nodes` is the file with the list of new database servers.

8. Verify the `ORACLE_HOME` directories have been added to the new servers.

```
# dcli -g /root/all_group -l root du -sm \
/u01/app/oracle/product/11.2.0/dbhome_1
```

3.10 Adding Database Instance to the New Servers

Before adding the database instances to the new servers, check the following:

- **Maximum file size:** If any data files have reached their maximum file size, then the `addInstance` command may fail with an ORA-00740 error. Oracle recommends you check that none of the files listed in `DBA_DATA_FILES` have reached their maximum size. Files that have reached their maximum should be corrected.
 - **Online redo logs:** If the online redo logs are kept in the directory specified by the `DB_RECOVERY_FILE_DEST` parameter, then ensure the space allocated is sufficient for the additional redo logs for the new instances being added. If necessary, then increase the size for the `DB_RECOVERY_FILE_DEST_SIZE` parameter.
 - **Total number of instances in the cluster:** Set the value of the initialization parameter `cluster_database_instances` in the SPFILE for each database to the total number of instances that will be in the cluster after adding the new servers.
 - **The HugePages settings are correctly configured on the new servers to match the existing servers.**
1. Use a command similar the following from an existing database server to add instances to the new servers. In the command, the instance, `dbm9`, is being added for server `dm02db01`.

```
dbca -silent -addInstance -gdbName dbm -nodeList dm02db01 -  
instanceName dbm9 \  
-sysDBAUsername sys
```

The command must be run for all servers and instances, substituting the server name and instance name, as appropriate.

Note:

If the command fails, then ensure any files that were created, such as redo log files, are cleaned up. The `deleteInstance` command does not clean log files or data files that were created by the `addInstance` command.

2. Add the `CLUSTER_INTERCONNECTS` parameter to each new instance.
 - a. Manually add the `CLUSTER_INTERCONNECTS` parameter to the SPFILE for each new database instance. The additions are similar to the existing entries, but are the RDMA Network Fabric addresses corresponding to the server that each instance runs on.
 - b. Restart the instance on each new server.
 - c. Verify the parameters were set correctly.

3.11 Returning the Rack to Service

Use the following procedure to ensure the new hardware is correctly configured and ready for use:

1. Verify the RDMA Network Fabric cables are connected and secure.
 - For RoCE, run the `verify_roce_cables.py` script, available from My Oracle Support.
 - For InfiniBand, run the `/opt/oracle.SupportTools/ibdiagtools/verify-topology` command.
2. Run the Oracle Exadata Database Machine HealthCheck utility using the steps described in My Oracle Support note 1070954.1.
3. Verify the instance additions using the following commands:

```
srvctl config database -d dbm
srvctl status database -d dbm
```

4. Check the cluster resources using the following command:

```
crsctl stat res -t
```

5. Ensure the original configuration summary report from the original cluster deployment is updated to include all servers. This document should include the calibrate and network verifications for the new rack, and the RDMA Network Fabric cable checks.
6. Conduct a power-off test, if possible. If the new Exadata Storage Servers cannot be powered off, then verify that the new database servers with the new instances can be powered off and powered on, and that all processes start automatically.

 **Note:**

Ensure the Oracle ASM disk rebalance process has completed for all disk groups. Use SQL*Plus to connect to an Oracle ASM instance and issue the following command:

```
SELECT * FROM gv$asm_operation;
```

No rows should be returned by the command.

7. Review the configuration settings, such as the following:
 - All parallelism settings
 - Backup configurations
 - Standby site, if any
 - Service configuration
 - Oracle Database File System (DBFS) configuration, and mount points on new servers (not required for X7 and later servers)
 - Installation of Oracle Enterprise HugePage Manager agents on new database servers
 - HugePages settings
8. Incorporate the new cell and database servers into Oracle Auto Service Request (ASR).
9. Update Oracle Enterprise Manager Cloud Control to include the new nodes.

Related Topics

- [Verifying InfiniBand Network Configuration](#)
- [About Oracle Auto Service Request](#)
- [Verify RoCE Cabling on Oracle Exadata Database Machine X8M-2 and X8M-8 Servers \(My Oracle Support Doc ID 2587717.1\)](#)
- [Oracle Exadata Database Machine EXAchk or HealthCheck \(My Oracle Support Doc ID 1070954.1\)](#)

4

Multi-Rack Cabling Tables for Oracle Exadata Rack X9M

This section contains multi-rack cabling tables for Oracle Exadata Rack X9M, which uses RoCE Network Fabric.

- [Understanding Multi-Rack Cabling for X9M Racks](#)
Up to 12 X9M racks can be cabled together without external RDMA Network Fabric switches.
- [Preparing for Multi-Rack Cabling with X9M Racks](#)
- [Two-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect two X9M racks together, both of which use RoCE Network Fabric.
- [Three-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect three X9M racks together using RoCE Network Fabric.
- [Four-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect four X9M racks together, all of which use RoCE Network Fabric.
- [Five-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect five (5) X9M racks together, all of which use RoCE Network Fabric.
- [Six-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect six (6) X9M racks together, all of which use RoCE Network Fabric.
- [Seven-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect seven (7) X9M racks together, all of which use RoCE Network Fabric.
- [Eight-Rack Cabling for X9M Racks](#)
This section provides the cabling details to connect eight (8) X9M racks together, all of which use RoCE Network Fabric.

4.1 Understanding Multi-Rack Cabling for X9M Racks

Up to 12 X9M racks can be cabled together without external RDMA Network Fabric switches.

For configurations containing eight or fewer racks, each rack must contain three RDMA Network Fabric switches.

Two of the switches, known as leaf switches, connect to all of the database servers and storage servers in the rack. The leaf switches are located as follows:

- At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2.
- At rack unit 21 (U21) and 23 (U23) in Exadata Database Machine X9M-8.

The third switch, known as the spine switch, interconnects the leaf switches in a multi-rack configuration. The spine switch is located at rack unit 1 (U1). If any rack does not include a spine switch, then order the expansion switch kit.

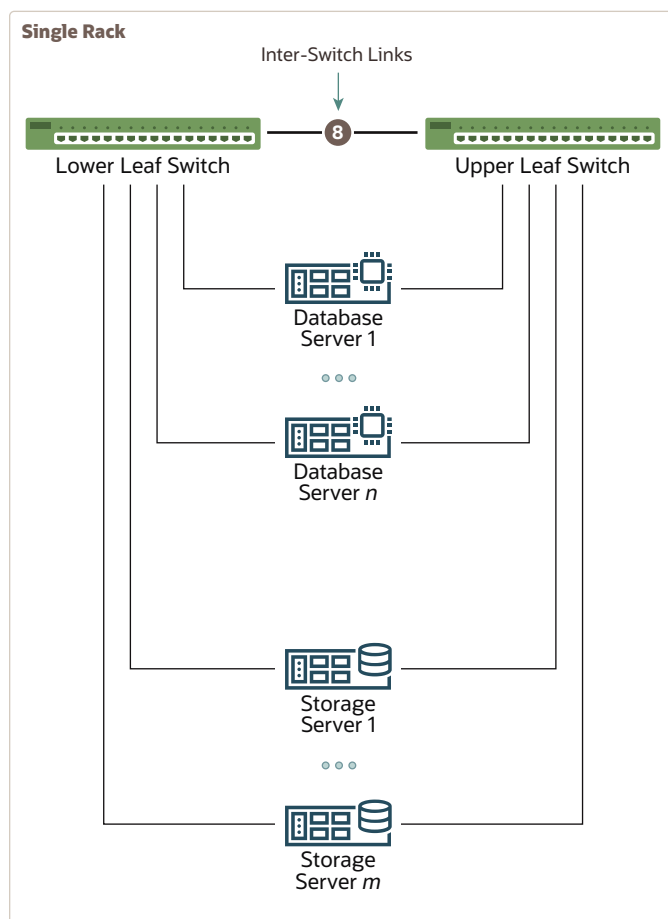
For configurations containing nine to 12 racks, no additional spine switches are required apart from those in the first eight racks.

The procedures in this section assume the racks are adjacent to each other, standard routing in raised floor is used, and there are no obstacles in the raised floor. If these assumptions are not correct for your environment, then longer cables may be required for the connections.

 **Note:**

By default, Oracle Exadata Database Machine X9M racks do not include spare cables or a third RoCE Network Fabric switch. To extend these racks, you must order the required cables and RoCE Network Fabric switch.

The following diagram shows the default RDMA Network Fabric architecture for a single-rack system. Each rack has two leaf switches, with eight connections between the leaf switches. The database servers and storage servers are each connected to both leaf switches. Each server contains a dual-port RDMA Network Fabric card, with port 1 connected to the lower leaf switch and port 2 connected to the upper leaf switch.

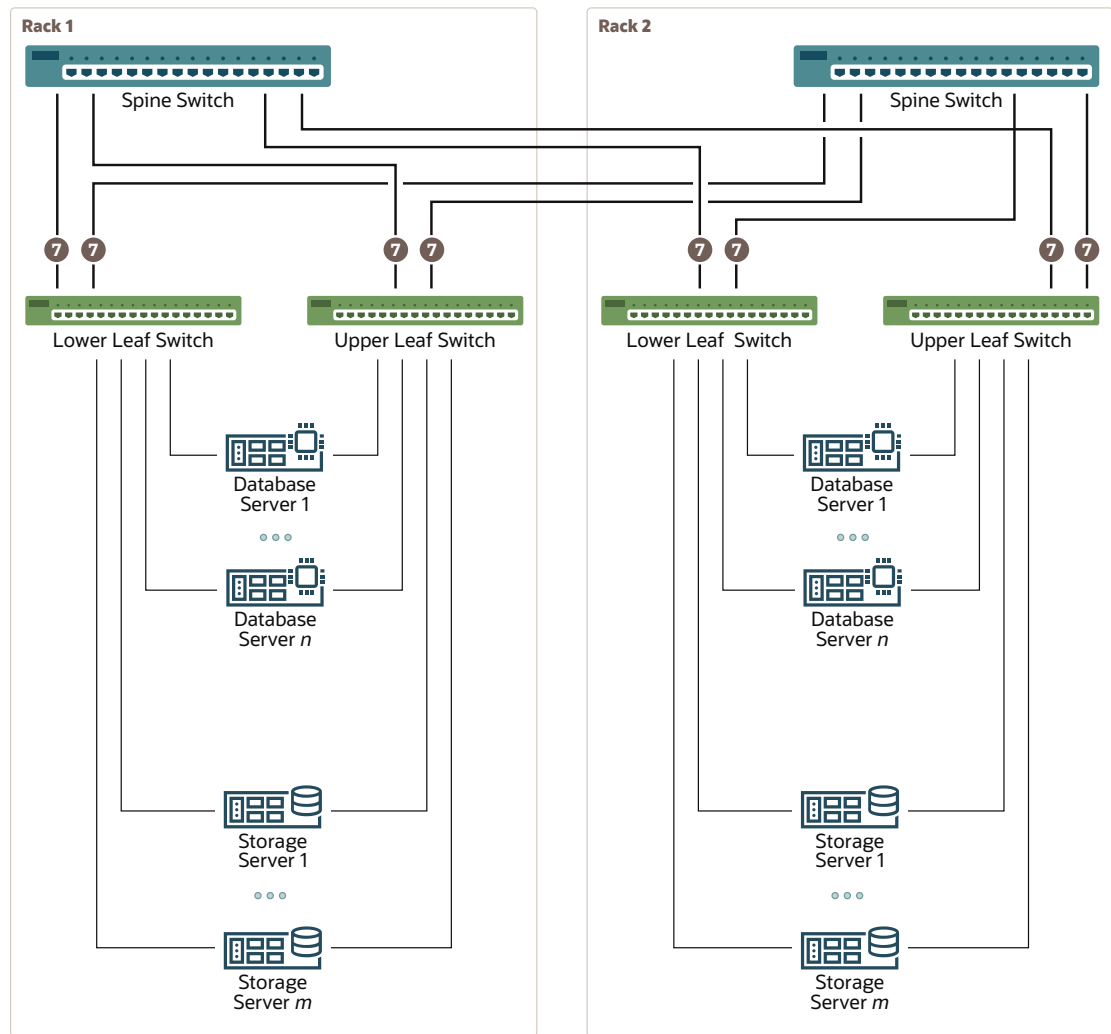


To connect up to eight X9M racks together, use the following general approach:

1. Remove the eight existing inter-switch connections between the leaf switches on each rack.
2. From each leaf switch, evenly distribute 14 connections to the spine switches in all of the interconnected racks.

The 14 connections use the 8 ports that were previously used for the inter-switch connections and 6 additional free ports on each leaf switch.

The resulting RoCE Network Fabric for a typical 2 rack system is illustrated in the following diagram:



As shown in the preceding diagram, every leaf switch has 7 connections to every spine switch. The leaf switches are not directly interconnected with other leaf switches, and the spine switches are not directly interconnected with each other.

As the number of racks increases, every leaf switch has 14 inter-switch connections that are evenly distributed across all of spine switches.

 **Note:**

For X9M-8 systems with three database servers and 11 storage servers only, the database servers and storage server require 23 leaf switch ports. Consequently, on such systems only, there are only 13 inter-switch links on each leaf switch.

After eight racks are connected, to connect additional X9M racks up to a total of 12 racks, use the following general approach:

1. Remove the eight existing inter-switch connections between the leaf switches on the additional rack.
2. From each leaf switch on the additional rack, evenly distribute 14 connections to the spine switches in the first 8 interconnected racks.

The 14 connections use the 8 ports that were previously used for the inter-switch connections and 6 additional free ports on each leaf switch.

4.2 Preparing for Multi-Rack Cabling with X9M Racks

Racks can be added together to increase system capacity and performance. When cabling racks together, note the following:

- The cable lengths shown in this document assume the racks are adjacent to each other. If the racks are not adjacent, or there are obstacles in the raised floor, or if you use overhead cabling, then longer cables may be required. For optical cables, the maximum supported cable length is 100 meters. For copper cables, the maximum supported cable length is 5 meters.
- Oracle recommends that the names for the servers include the rack unit number. This helps identify the server during diagnostics.
- When completing Oracle Exadata Deployment Assistant (OEDA) for the additional rack, you are prompted for SCAN addresses. However, these SCAN addresses are not used because the SCAN address from the original rack are used. Manually remove the new SCAN addresses from the generated installation files.
- The software owner account names and group names, as well as their identifiers, must match the names and identifiers of the original rack.
- If the additional grid disks are used with existing disk groups, then ensure the grid disk sizes for the new rack are the same as the original rack.
- ***For multi-rack configurations containing eight or fewer racks, a spine switch must exist in each rack in order to interconnect the RoCE Network Fabric.***

Perform the following tasks before cabling racks together:

1. Determine the number of racks that will be cabled together.
2. Order the parts needed to connect the racks.

To extend Exadata Database Machine racks with RoCE Network Fabric, for each rack being added you must order extra cables, transceivers for longer cables, and a RoCE Network Fabric spine switch, if one is required.

When connecting four or more racks, or if you need longer cables for your environment, you must purchase additional 10 meter fiber cables with two QSFP28 SR transceivers to connect each end. The QSFP28 SR transceivers are needed for fiber cables over 5 meters in length.

For multi-rack configurations containing eight or fewer racks, the following table outlines the cables needed to interconnect the racks:

Number of Racks	3 Meter Cables Needed	5 Meter Cables Needed	10 Meter Cables Needed
2	28	28	0
3	28	56	0
4	28	71	13
5	28	78	34
6	28	87	53
7	28	88	80
8	28	91	105

3. Gather the following tools:
 - Pen
 - Note pad
 - Flashlight
 - Labels for cables and rack
 - Cable ties (Oracle recommends hook and loop straps)
4. Determine a naming method for the rack prefixes. For example, if the original rack has the prefix `dbm01`, then use the prefix `dbm02` for the second rack, the prefix `dbm03` for the third rack, and so on.
5. Verify that the racks have unique host names and IP addresses. All servers interconnected in the racks must have unique names and IP addresses.

Server name and IP address conventions may differ in the following cases:

- **Initial installation of all Exadata Database Machine Racks:** System address assignments and host names should be complete.
 - **New Exadata Database Machine Racks are added to an existing cluster:** The new rack configuration should require unique host names and IP addresses for the new Exadata Database Machine. The IP addresses on the same subnet cannot conflict with the existing systems.
 - **Two existing Exadata Database Machine Racks are clustered together:** You can assign host names and IP addresses only if Exadata Database Machines are already assigned unique host names and IP addresses, or the entire cluster must be reconfigured. The machines must be on the same subnet and not have conflicting IP addresses.
6. Ensure the IP addresses for the new servers are in the same subnet, and do not overlap with the currently-installed servers.
 7. Check that the firmware on the original switches is at the same level as the new switches by using the `sh ver` command.

It is highly recommended, though not mandatory, to use the same firmware version on all of the switches. If the firmware is not at the same level, you can apply a firmware patch to bring the switches up to the same firmware level.

The remaining topics in this section provide detailed cabling information for cabling up to 8 racks together.

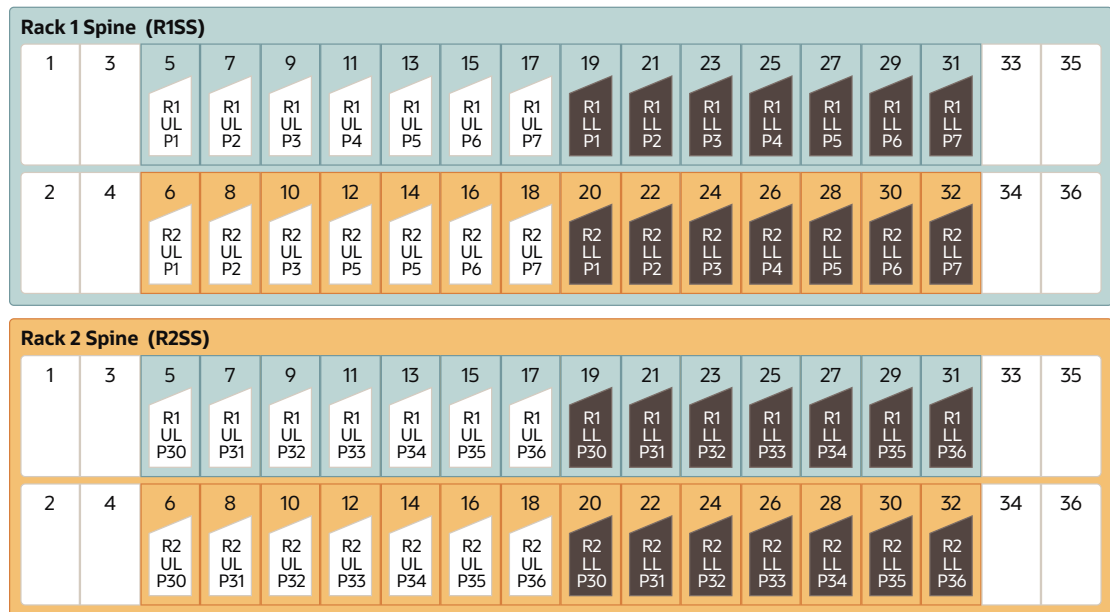
4.3 Two-Rack Cabling for X9M Racks

This section provides the cabling details to connect two X9M racks together, both of which use RoCE Network Fabric.

Note:

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the spine switches when cabling two racks together:



The following tables contain details for all of the RoCE Network Fabric cabling connections in a two-rack system.

Table 4-1 Leaf Switch Connections for the First Rack in a Two-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7 R1UL-P3 to R1SS-P9 R1UL-P4 to R1SS-P11 R1UL-P5 to R1SS-P13 R1UL-P6 to R1SS-P15 R1UL-P7 to R1SS-P17	3 meters
R1UL to R2SS	R1UL-P30 to R2SS-P5 R1UL-P31 to R2SS-P7 R1UL-P32 to R2SS-P9 R1UL-P33 to R2SS-P11 R1UL-P34 to R2SS-P13 R1UL-P35 to R2SS-P15 R1UL-P36 to R2SS-P17	5 meters
R1LL to R1SS	R1LL-P1 to R1SS-P19 R1LL-P2 to R1SS-P21 R1LL-P3 to R1SS-P23 R1LL-P4 to R1SS-P25 R1LL-P5 to R1SS-P27 R1LL-P6 to R1SS-P29 R1LL-P7 to R1SS-P31	3 meters

Table 4-1 (Cont.) Leaf Switch Connections for the First Rack in a Two-Rack System

Switch Pair	Connections	Cable Length
R1LL to R2SS	R1LL-P30 to R2SS-P19 R1LL-P31 to R2SS-P21 R1LL-P32 to R2SS-P23 R1LL-P33 to R2SS-P25 R1LL-P34 to R2SS-P27 R1LL-P35 to R2SS-P29 R1LL-P36 to R2SS-P31	5 meters

Table 4-2 Leaf Switch Connections for the Second Rack in a Two-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8 R2UL-P3 to R1SS-P10 R2UL-P4 to R1SS-P12 R2UL-P5 to R1SS-P14 R2UL-P6 to R1SS-P16 R2UL-P7 to R1SS-P18	5 meters
R2UL to R2SS	R2UL-P30 to R2SS-P6 R2UL-P31 to R2SS-P8 R2UL-P32 to R2SS-P10 R2UL-P33 to R2SS-P12 R2UL-P34 to R2SS-P14 R2UL-P35 to R2SS-P16 R2UL-P36 to R2SS-P18	3 meters
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22 R2LL-P3 to R1SS-P24 R2LL-P4 to R1SS-P26 R2LL-P5 to R1SS-P28 R2LL-P6 to R1SS-P30 R2LL-P7 to R1SS-P32	5 meters
R2LL to R2SS	R2LL-P30 to R2SS-P20 R2LL-P31 to R2SS-P22 R2LL-P32 to R2SS-P24 R2LL-P33 to R2SS-P26 R2LL-P34 to R2SS-P28 R2LL-P35 to R2SS-P30 R2LL-P36 to R2SS-P32	3 meters

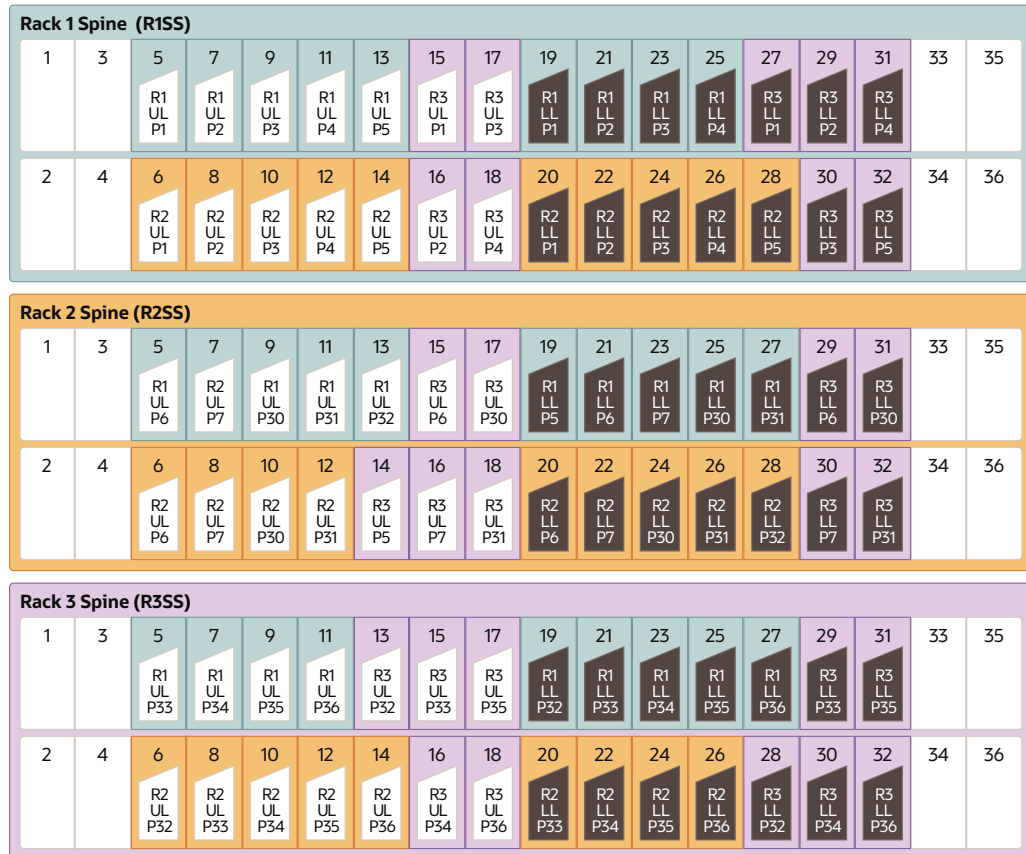
4.4 Three-Rack Cabling for X9M Racks

This section provides the cabling details to connect three X9M racks together using RoCE Network Fabric.

Note:

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the spine switches when cabling three racks together:



The following tables contain details for all of the RoCE Network Fabric cabling connections in a three-rack system.

Table 4-3 Leaf Switch Connections for the First Rack in a Three-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7 R1UL-P3 to R1SS-P9 R1UL-P4 to R1SS-P11 R1UL-P5 to R1SS-P13	3 meters
R1UL to R2SS	R1UL-P6 to R2SS-P5 R1UL-P7 to R2SS-P7 R1UL-P30 to R2SS-P9 R1UL-P31 to R2SS-P11 R1UL-P32 to R2SS-P13	5 meters
R1UL to R3SS	R1UL-P33 to R3SS-P5 R1UL-P34 to R3SS-P7 R1UL-P35 to R3SS-P9 R1UL-P36 to R3SS-P11	5 meters

Table 4-3 (Cont.) Leaf Switch Connections for the First Rack in a Three-Rack System

Switch Pair	Connections	Cable Length
R1LL to R1SS	R1LL-P1 to R1SS-P19 R1LL-P2 to R1SS-P21 R1LL-P3 to R1SS-P23 R1LL-P4 to R1SS-P25	3 meters
R1LL to R2SS	R1LL-P5 to R2SS-P19 R1LL-P6 to R2SS-P21 R1LL-P7 to R2SS-P23 R1LL-P30 to R2SS-P25 R1LL-P31 to R2SS-P27	5 meters
R1LL to R3SS	R1LL-P32 to R3SS-P19 R1LL-P33 to R3SS-P21 R1LL-P34 to R3SS-P23 R1LL-P35 to R3SS-P25 R1LL-P36 to R3SS-P27	5 meters

Table 4-4 Leaf Switch Connections for the Second Rack in a Three-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8 R2UL-P3 to R1SS-P10 R2UL-P4 to R1SS-P12 R2UL-P5 to R1SS-P14	5 meters
R2UL to R2SS	R2UL-P6 to R2SS-P6 R2UL-P7 to R2SS-P8 R2UL-P30 to R2SS-P10 R2UL-P31 to R2SS-P12	3 meters
R2UL to R3SS	R2UL-P32 to R3SS-P6 R2UL-P33 to R3SS-P8 R2UL-P34 to R3SS-P10 R2UL-P35 to R3SS-P12 R2UL-P36 to R3SS-P14	5 meters
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22 R2LL-P3 to R1SS-P24 R2LL-P4 to R1SS-P26 R2LL-P5 to R1SS-P28	5 meters

Table 4-4 (Cont.) Leaf Switch Connections for the Second Rack in a Three-Rack System

Switch Pair	Connections	Cable Length
R2LL to R2SS	R2LL-P6 to R2SS-P20 R2LL-P7 to R2SS-P22 R2LL-P30 to R2SS-P24 R2LL-P31 to R2SS-P26 R2LL-P32 to R2SS-P28	3 meters
R2LL to R3	R2LL-P33 to R3SS-P20 R2LL-P34 to R3SS-P22 R2LL-P35 to R3SS-P24 R2LL-P36 to R3SS-P26	5 meters

Table 4-5 Leaf Switch Connections for the Third Rack in a Three-Rack System

Switch Pair	Connections	Cable Length
R3UL to R1SS	R3UL-P1 to R1SS-P15 R3UL-P2 to R1SS-P16 R3UL-P3 to R1SS-P17 R3UL-P4 to R1SS-P18	5 meters
R3UL to R2SS	R3UL-P5 to R2SS-P14 R3UL-P6 to R2SS-P15 R3UL-P7 to R2SS-P16 R3UL-P30 to R2SS-P17 R3UL-P31 to R2SS-P18	5 meters
R3UL to R3SS	R3UL-P32 to R3SS-P13 R3UL-P33 to R3SS-P15 R3UL-P34 to R3SS-P16 R3UL-P35 to R3SS-P17 R3UL-P36 to R3SS-P18	3 meters
R3LL to R1SS	R3LL-P1 to R1SS-P27 R3LL-P2 to R1SS-P29 R3LL-P3 to R1SS-P30 R3LL-P4 to R1SS-P31 R3LL-P5 to R1SS-P32	5 meters
R3LL to R2SS	R3LL-P6 to R2SS-P29 R3LL-P7 to R2SS-P30 R3LL-P30 to R2SS-P31 R3LL-P31 to R2SS-P32	5 meters
R3LL to R3SS	R3LL-P32 to R3SS-P27 R3LL-P33 to R3SS-P29 R3LL-P34 to R3SS-P30 R3LL-P35 to R3SS-P31 R3LL-P36 to R3SS-P32	3 meters

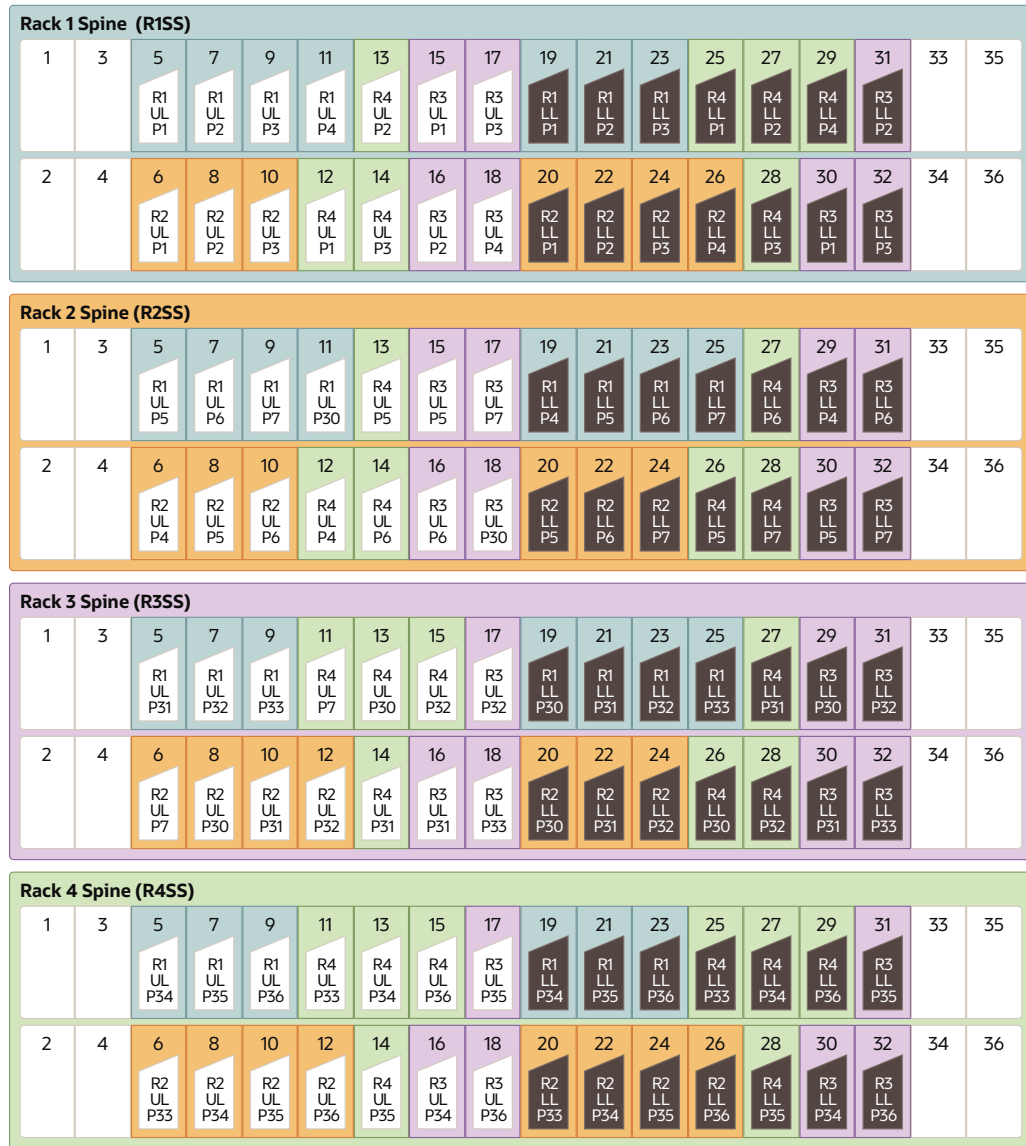
4.5 Four-Rack Cabling for X9M Racks

This section provides the cabling details to connect four X9M racks together, all of which use RoCE Network Fabric.

Note:

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the spine switches when cabling four racks together:



The following tables contain details for all of the RoCE Network Fabric cabling connections in a four-rack system.

Table 4-6 Leaf Switch Connections for the First Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7 R1UL-P3 to R1SS-P9 R1UL-P4 to R1SS-P11	3 meters
R1UL to R2SS	R1UL-P5 to R2SS-P5 R1UL-P6 to R2SS-P7 R1UL-P7 to R2SS-P9 R1UL-P30 to R2SS-P11	5 meters

Table 4-6 (Cont.) Leaf Switch Connections for the First Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R1UL to R3SS	R1UL-P31 to R3SS-P5 R1UL-P32 to R3SS-P7 R1UL-P33 to R3SS-P9	5 meters
R1UL to R4SS	R1UL-P34 to R4SS-P5 R1UL-P35 to R4SS-P7 R1UL-P36 to R4SS-P9	10 meters
R1LL to R1SS	R1LL-P1 to R1SS-P19 R1LL-P2 to R1SS-P21 R1LL-P3 to R1SS-P23	3 meters
R1LL to R2SS	R1LL-P4 to R2SS-P19 R1LL-P5 to R2SS-P21 R1LL-P6 to R2SS-P23 R1LL-P7 to R2SS-P25	5 meters
R1LL to R3SS	R1LL-P30 to R3SS-P19 R1LL-P31 to R3SS-P21 R1LL-P32 to R3SS-P23 R1LL-P33 to R3SS-P25	5 meters
R1LL to R4SS	R1LL-P34 to R4SS-P19 R1LL-P35 to R4SS-P21 R1LL-P36 to R4SS-P23	10 meters

Table 4-7 Leaf Switch Connections for the Second Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8 R2UL-P3 to R1SS-P10	5 meters
R2UL to R2SS	R2UL-P4 to R2SS-P6 R2UL-P5 to R2SS-P8 R2UL-P6 to R2SS-P10	3 meters
R2UL to R3SS	R2UL-P7 to R3SS-P6 R2UL-P30 to R3SS-P8 R2UL-P31 to R3SS-P10 R2UL-P32 to R3SS-P12	5 meters
R2UL to R4SS	R2UL-P33 to R4SS-P6 R2UL-P34 to R4SS-P8 R2UL-P35 to R4SS-P10 R2UL-P36 to R4SS-P12	5 meters

Table 4-7 (Cont.) Leaf Switch Connections for the Second Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22 R2LL-P3 to R1SS-P24 R2LL-P4 to R1SS-P26	5 meters
R2LL to R2SS	R2LL-P5 to R2SS-P20 R2LL-P6 to R2SS-P22 R2LL-P7 to R2SS-P24	3 meters
R2LL to R3SS	R2LL-P30 to R3SS-P20 R2LL-P31 to R3SS-P22 R2LL-P32 to R3SS-P24	5 meters
R2LL to R4SS	R2LL-P33 to R4SS-P20 R2LL-P34 to R4SS-P22 R2LL-P35 to R4SS-P24 R2LL-P36 to R4SS-P26	5 meters

Table 4-8 Leaf Switch Connections for the Third Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R3UL to R1SS	R3UL-P1 to R1SS-P15 R3UL-P2 to R1SS-P16 R3UL-P3 to R1SS-P17 R3UL-P4 to R1SS-P18	5 meters
R3UL to R2SS	R3UL-P5 to R2SS-P15 R3UL-P6 to R2SS-P16 R3UL-P7 to R2SS-P17 R3UL-P30 to R2SS-P18	5 meters
R3UL to R3SS	R3UL-P31 to R3SS-P16 R3UL-P32 to R3SS-P17 R3UL-P33 to R3SS-P18	3 meters
R3UL to R4SS	R3UL-P34 to R4SS-P16 R3UL-P35 to R4SS-P17 R3UL-P36 to R4SS-P18	5 meters
R3LL to R1SS	R3LL-P1 to R1SS-P30 R3LL-P2 to R1SS-P31 R3LL-P3 to R1SS-P32	5 meters
R3LL to R2SS	R3LL-P4 to R2SS-P29 R3LL-P5 to R2SS-P30 R3LL-P6 to R2SS-P31 R3LL-P7 to R2SS-P32	5 meters

Table 4-8 (Cont.) Leaf Switch Connections for the Third Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R3LL to R3SS	R3LL-P30 to R3SS-P29 R3LL-P31 to R3SS-P30 R3LL-P32 to R3SS-P31 R3LL-P33 to R3SS-P32	3 meters
R3LL to R4SS	R3LL-P34 to R4SS-P30 R3LL-P35 to R4SS-P31 R3LL-P36 to R4SS-P32	5 meters

Table 4-9 Leaf Switch Connections for the Fourth Rack in a Four-Rack System

Switch Pair	Connections	Cable Length
R4UL to R1SS	R4UL-P1 to R1SS-P12 R4UL-P2 to R1SS-P13 R4UL-P3 to R1SS-P14	10 meters
R4UL to R2SS	R4UL-P4 to R2SS-P12 R4UL-P5 to R2SS-P13 R4UL-P6 to R2SS-P14	5 meters
R4UL to R3SS	R4UL-P7 to R3SS-P11 R4UL-P30 to R3SS-P13 R4UL-P31 to R3SS-P14 R4UL-P32 to R3SS-P15	5 meters
R4UL to R4SS	R4UL-P33 to R4SS-P11 R4UL-P34 to R4SS-P13 R4UL-P35 to R4SS-P14 R4UL-P36 to R4SS-P15	3 meters
R4LL to R1SS	R4LL-P1 to R1SS-P25 R4LL-P2 to R1SS-P27 R4LL-P3 to R1SS-P28 R4LL-P4 to R1SS-P29	10 meters
R4LL to R2SS	R4LL-P5 to R2SS-P26 R4LL-P6 to R2SS-P27 R4LL-P7 to R2SS-P28	5 meters
R4LL to R3SS	R4LL-P30 to R3SS-P26 R4LL-P31 to R3SS-P27 R4LL-P32 to R3SS-P28	5 meters
R4LL to R4SS	R4LL-P33 to R4SS-P25 R4LL-P34 to R4SS-P27 R4LL-P35 to R4SS-P28 R4LL-P36 to R4SS-P29	3 meters

4.6 Five-Rack Cabling for X9M Racks

This section provides the cabling details to connect five (5) X9M racks together, all of which use RoCE Network Fabric.

Note:

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the spine switches when cabling five racks together:

Rack 1 Spine (R1SS)

1	3	5 R1 UL P1	7 R1 UL P2	9 R1 UL P3	11 R5 UL P1	13 R4 UL P1	15 R5 UL P3	17 R3 UL P2	19 R1 LL P1	21 R1 LL P2	23 R5 LL P1	25 R5 LL P2	27 R4 LL P1	29 R4 LL P3	31 R3 LL P2	33	35
2	4	6 R2 UL P1	8 R2 UL P2	10 R2 UL P3	12 R5 UL P2	14 R4 UL P2	16 R3 UL P1	18 R3 UL P3	20 R2 LL P1	22 R2 LL P2	24 R2 LL P3	26 R5 LL P3	28 R4 LL P2	30 R3 LL P1	32 R3 LL P3	34	36

Rack 2 Spine (R2SS)

1	3	5 R1 UL P4	7 R1 UL P5	9 R1 UL P6	11 R5 UL P5	13 R4 UL P4	15 R5 UL P6	17 R3 UL P5	19 R1 LL P3	21 R1 LL P4	23 R1 LL P5	25 R5 LL P4	27 R4 LL P4	29 R5 LL P6	31 R3 LL P5	33	35
2	4	6 R2 UL P4	8 R2 UL P5	10 R5 UL P4	12 R4 UL P3	14 R4 UL P5	16 R3 UL P4	18 R3 UL P6	20 R2 LL P4	22 R2 LL P5	24 R2 LL P6	26 R5 LL P5	28 R4 LL P5	30 R3 LL P4	32 R3 LL P6	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P7	7 R1 UL P30	9 R1 UL P31	11 R5 UL P7	13 R4 UL P6	15 R4 UL P30	17 R3 UL P30	19 R1 LL P6	21 R1 LL P7	23 R1 LL P30	25 R5 LL P30	27 R4 LL P7	29 R5 LL P31	31 R3 LL P30	33	35
2	4	6 R2 UL P6	8 R2 UL P7	10 R2 UL P30	12 R5 UL P30	14 R4 UL P7	16 R3 UL P7	18 R3 UL P31	20 R2 LL P7	22 R2 LL P30	24 R5 LL P7	26 R4 LL P6	28 R4 LL P30	30 R3 LL P7	32 R3 LL P31	34	36

Rack 4 Spine (R4SS)

1	3	5 R1 UL P32	7 R1 UL P33	9 R1 UL P34	11 R5 UL P31	13 R4 UL P31	15 R4 UL P33	17 R3 UL P32	19 R1 LL P31	21 R1 LL P32	23 R1 LL P33	25 R5 LL P32	27 R4 LL P31	29 R4 LL P33	31 R3 LL P33	33	35
2	4	6 R2 UL P31	8 R2 UL P32	10 R2 UL P33	12 R5 UL P32	14 R4 UL P32	16 R5 UL P33	18 R3 UL P33	20 R2 LL P31	22 R2 LL P32	24 R2 LL P33	26 R5 LL P32	28 R4 LL P32	30 R3 LL P32	32 R3 LL P34	34	36

Rack 5 Spine (R5SS)

1	3	5 R1 UL P35	7 R1 UL P36	9 R5 UL P34	11 R5 UL P35	13 R4 UL P34	15 R4 UL P36	17 R3 UL P35	19 R1 LL P34	21 R1 LL P35	23 R1 LL P36	25 R4 LL P34	27 R4 LL P35	29 R4 LL P36	31 R3 LL P35	33	35
2	4	6 R2 UL P34	8 R2 UL P35	10 R2 UL P36	12 R5 UL P36	14 R4 UL P35	16 R3 UL P34	18 R3 UL P36	20 R2 LL P34	22 R2 LL P35	24 R2 LL P36	26 R5 LL P34	28 R5 LL P35	30 R5 LL P36	32 R3 LL P36	34	36

The following tables contain details for all of the RoCE Network Fabric cabling connections in a five-rack system.

Table 4-10 Leaf Switch Connections for the First Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7 R1UL-P3 to R1SS-P9	3 meters
R1UL to R2SS	R1UL-P4 to R2SS-P5 R1UL-P5 to R2SS-P7 R1UL-P6 to R2SS-P9	5 meters
R1UL to R3SS	R1UL-P7 to R3SS-P5 R1UL-P30 to R3SS-P7 R1UL-P31 to R3SS-P9	5 meters
R1UL to R4SS	R1UL-P32 to R4SS-P5 R1UL-P33 to R4SS-P7 R1UL-P34 to R4SS-P9	10 meters
R1UL to R5SS	R1UL-P35 to R5SS-P5 R1UL-P36 to R5SS-P7	10 meters
R1LL to R1SS	R1LL-P1 to R1SS-P19 R1LL-P2 to R1SS-P21	3 meters
R1LL to R2SS	R1LL-P3 to R2SS-P19 R1LL-P4 to R2SS-P21 R1LL-P5 to R2SS-P23	5 meters
R1LL to R3SS	R1LL-P6 to R3SS-P19 R1LL-P7 to R3SS-P21 R1LL-P30 to R3SS-P23	5 meters
R1LL to R4SS	R1LL-P31 to R4SS-P19 R1LL-P32 to R4SS-P21 R1LL-P33 to R4SS-P23	10 meters
R1LL to R5SS	R1LL-P34 to R5SS-P19 R1LL-P35 to R5SS-P21 R1LL-P36 to R5SS-P23	10 meters

Table 4-11 Leaf Switch Connections for the Second Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8 R2UL-P3 to R1SS-P10	5 meters
R2UL to R2SS	R2UL-P4 to R2SS-P6 R2UL-P5 to R2SS-P8	3 meters
R2UL to R3SS	R2UL-P6 to R3SS-P6 R2UL-P7 to R3SS-P8 R2UL-P30 to R3SS-P10	5 meters

Table 4-11 (Cont.) Leaf Switch Connections for the Second Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R2UL to R4SS	R2UL-P31 to R4SS-P6 R2UL-P32 to R4SS-P8 R2UL-P33 to R4SS-P10	5 meters
R2UL to R5SS	R2UL-P34 to R5SS-P6 R2UL-P35 to R5SS-P8 R2UL-P36 to R5SS-P10	10 meters
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22 R2LL-P3 to R1SS-P24	5 meters
R2LL to R2SS	R2LL-P4 to R2SS-P20 R2LL-P5 to R2SS-P22 R2LL-P6 to R2SS-P24	3 meters
R2LL to R3SS	R2LL-P7 to R3SS-P20 R2LL-P30 to R3SS-P22	5 meters
R2LL to R4SS	R2LL-P31 to R4SS-P20 R2LL-P32 to R4SS-P22 R2LL-P33 to R4SS-P24	5 meters
R2LL to R5SS	R2LL-P34 to R5SS-P20 R2LL-P35 to R5SS-P22 R2LL-P36 to R5SS-P24	10 meters

Table 4-12 Leaf Switch Connections for the Third Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R3UL to R1SS	R3UL-P1 to R1SS-P16 R3UL-P2 to R1SS-P17 R3UL-P3 to R1SS-P18	5 meters
R3UL to R2SS	R3UL-P4 to R2SS-P16 R3UL-P5 to R2SS-P17 R3UL-P6 to R2SS-P18	5 meters
R3UL to R3SS	R3UL-P7 to R3SS-P16 R3UL-P30 to R3SS-P17 R3UL-P31 to R3SS-P18	3 meters
R3UL to R4SS	R3UL-P32 to R4SS-P17 R3UL-P33 to R4SS-P18	5 meters
R3UL to R5SS	R3UL-P34 to R5SS-P16 R3UL-P35 to R5SS-P17 R3UL-P36 to R5SS-P18	5 meters

Table 4-12 (Cont.) Leaf Switch Connections for the Third Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R3LL to R1SS	R3LL-P1 to R1SS-P30 R3LL-P2 to R1SS-P31 R3LL-P3 to R1SS-P32	5 meters
R3LL to R2SS	R3LL-P4 to R2SS-P30 R3LL-P5 to R2SS-P31 R3LL-P6 to R2SS-P32	5 meters
R3LL to R3SS	R3LL-P7 to R3SS-P30 R3LL-P30 to R3SS-P31 R3LL-P31 to R3SS-P32	3 meters
R3LL to R4SS	R3LL-P32 to R4SS-P30 R3LL-P33 to R4SS-P31 R3LL-P34 to R4SS-P32	5 meters
R3LL to R5SS	R3LL-P35 to R5SS-P31 R3LL-P36 to R5SS-P32	5 meters

Table 4-13 Leaf Switch Connections for the Fourth Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R4UL to R1SS	R4UL-P1 to R1SS-P13 R4UL-P2 to R1SS-P14	10 meters
R4UL to R2SS	R4UL-P3 to R2SS-P12 R4UL-P4 to R2SS-P13 R4UL-P5 to R2SS-P14	5 meters
R4UL to R3SS	R4UL-P6 to R3SS-P13 R4UL-P7 to R3SS-P14 R4UL-P30 to R3SS-P15	5 meters
R4UL to R4SS	R4UL-P31 to R4SS-P13 R4UL-P32 to R4SS-P14 R4UL-P33 to R4SS-P15	3 meters
R4UL to R5SS	R4UL-P34 to R5SS-P13 R4UL-P35 to R5SS-P14 R4UL-P36 to R5SS-P15	5 meters
R4LL to R1SS	R4LL-P1 to R1SS-P27 R4LL-P2 to R1SS-P28 R4LL-P3 to R1SS-P29	10 meters
R4LL to R2SS	R4LL-P4 to R2SS-P27 R4LL-P5 to R2SS-P28	5 meters
R4LL to R3SS	R4LL-P6 to R3SS-P26 R4LL-P7 to R3SS-P27 R4LL-P30 to R3SS-P28	5 meters

Table 4-13 (Cont.) Leaf Switch Connections for the Fourth Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R4LL to R4SS	R4LL-P31 to R4SS-P27 R4LL-P32 to R4SS-P28 R4LL-P33 to R4SS-P29	3 meters
R4LL to R5SS	R4LL-P34 to R5SS-P25 R4LL-P35 to R5SS-P27 R4LL-P36 to R5SS-P29	5 meters

Table 4-14 Leaf Switch Connections for the Fifth Rack in a Five-Rack System

Switch Pair	Connections	Cable Length
R5UL to R1SS	R5UL-P1 to R1SS-P11 R5UL-P2 to R1SS-P12 R5UL-P3 to R1SS-P15	10 meters
R5UL to R2SS	R5UL-P4 to R2SS-P10 R5UL-P5 to R2SS-P11 R5UL-P6 to R2SS-P15	10 meters
R5UL to R3SS	R5UL-P7 to R3SS-P11 R5UL-P30 to R3SS-P12	5 meters
R5UL to R4SS	R5UL-P31 to R4SS-P11 R5UL-P32 to R4SS-P12 R5UL-P33 to R4SS-P16	5 meters
R5UL to R5SS	R5UL-P34 to R5SS-P9 R5UL-P35 to R5SS-P11 R5UL-P36 to R5SS-P12	3 meters
R5LL to R1SS	R5LL-P1 to R1SS-P23 R5LL-P2 to R1SS-P25 R5LL-P3 to R1SS-P26	10 meters
R5LL to R2SS	R5LL-P4 to R2SS-P25 R5LL-P5 to R2SS-P26 R5LL-P6 to R2SS-P29	10 meters
R5LL to R3SS	R5LL-P7 to R3SS-P24 R5LL-P30 to R3SS-P25 R5LL-P31 to R3SS-P29	5 meters
R5LL to R4SS	R5LL-P32 to R4SS-P25 R5LL-P33 to R4SS-P26	5 meters
R5LL to R5SS	R5LL-P34 to R5SS-P26 R5LL-P35 to R5SS-P28 R5LL-P36 to R5SS-P30	3 meters

4.7 Six-Rack Cabling for X9M Racks

This section provides the cabling details to connect six (6) X9M racks together, all of which use RoCE Network Fabric.

Note:

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following diagrams show the cable connections for the spine switches when cabling six racks together:

Rack 1 Spine (R1SS)

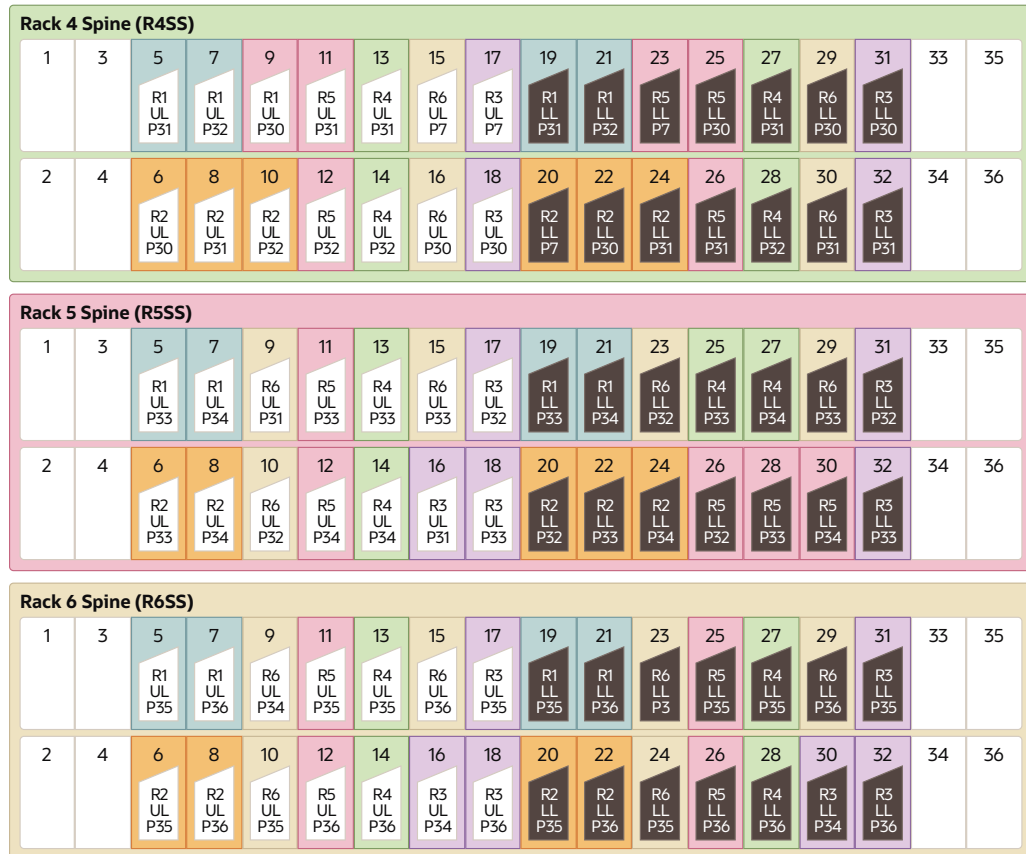
1	3	5 R1 UL P1	7 R1 UL P2	9 R1 UL P3	11 R5 UL P1	13 R4 UL P1	15 R4 UL P3	17 R3 UL P1	19 R1 LL P1	21 R1 LL P2	23 R6 LL P1	25 R5 LL P1	27 R4 LL P1	29 R6 LL P3	31 R3 LL P2	33	35
2	4	6 R2 UL P1	8 R2 UL P2	10 R6 UL P1	12 R5 UL P2	14 R4 UL P2	16 R6 UL P2	18 R3 UL P2	20 R2 LL P1	22 R2 LL P2	24 R6 LL P2	26 R5 LL P2	28 R4 LL P2	30 R3 LL P1	32 R3 LL P3	34	36

Rack 2 Spine (R2SS)

1	3	5 R1 UL P4	7 R1 UL P5	9 R1 UL P6	11 R5 UL P4	13 R4 UL P5	15 R6 UL P3	17 R3 UL P3	19 R1 LL P3	21 R1 LL P4	23 R1 LL P5	25 R5 LL P3	27 R4 LL P4	29 R5 LL P4	31 R3 LL P4	33	35
2	4	6 R2 UL P3	8 R2 UL P4	10 R5 UL P3	12 R4 UL P4	14 R4 UL P6	16 R6 UL P4	18 R3 UL P4	20 R2 LL P3	22 R2 LL P4	24 R6 LL P4	26 R4 LL P3	28 R4 LL P5	30 R6 LL P5	32 R3 LL P5	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P7	7 R1 UL P30	9 R5 UL P5	11 R5 UL P6	13 R4 UL P7	15 R6 UL P5	17 R3 UL P5	19 R1 LL P6	21 R1 LL P7	23 R1 LL P30	25 R5 LL P6	27 R4 LL P7	29 R6 LL P6	31 R3 LL P6	33	35
2	4	6 R2 UL P5	8 R2 UL P6	10 R2 UL P7	12 R5 UL P7	14 R4 UL P30	16 R6 UL P6	18 R3 UL P6	20 R2 LL P5	22 R2 LL P6	24 R5 LL P5	26 R4 LL P6	28 R4 LL P30	30 R6 LL P7	32 R3 LL P7	34	36



The following tables contain details for all of the RoCE Network Fabric cabling connections in a six-rack system.

Table 4-15 Leaf Switch Connections for the First Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7 R1UL-P3 to R1SS-P9	3 meters
R1UL to R2SS	R1UL-P4 to R2SS-P5 R1UL-P5 to R2SS-P7 R1UL-P6 to R2SS-P9	5 meters
R1UL to R3SS	R1UL-P7 to R3SS-P5 R1UL-P30 to R3SS-P7	5 meters
R1UL to R4SS	R1UL-P31 to R4SS-P5 R1UL-P32 to R4SS-P7	10 meters
R1UL to R5SS	R1UL-P33 to R5SS-P5 R1UL-P34 to R5SS-P7	10 meters
R1UL to R6SS	R1UL-P35 to R6SS-P5 R1UL-P36 to R6SS-P7	10 meters
R1LL to R1SS	R1LL-P1 to R1SS-P19 R1LL-P2 to R1SS-P21	3 meters

Table 4-15 (Cont.) Leaf Switch Connections for the First Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R1LL to R2SS	R1LL-P3 to R2SS-P19 R1LL-P4 to R2SS-P21 R1LL-P5 to R2SS-P23	5 meters
R1LL to R3SS	R1LL-P6 to R3SS-P19 R1LL-P7 to R3SS-P21 R1LL-P30 to R3SS-P23	5 meters
R1LL to R4SS	R1LL-P31 to R4SS-P19 R1LL-P32 to R4SS-P21	10 meters
R1LL to R5SS	R1LL-P33 to R5SS-P19 R1LL-P34 to R5SS-P21	10 meters
R1LL to R6SS	R1LL-P35 to R6SS-P19 R1LL-P36 to R6SS-P21	10 meters

Table 4-16 Leaf Switch Connections for the Second Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8	5 meters
R2UL to R2SS	R2UL-P3 to R2SS-P6 R2UL-P4 to R2SS-P8	3 meters
R2UL to R3SS	R2UL-P5 to R3SS-P6 R2UL-P6 to R3SS-P8 R2UL-P7 to R3SS-P10	5 meters
R2UL to R4SS	R2UL-P30 to R4SS-P6 R2UL-P31 to R4SS-P8 R2UL-P32 to R4SS-P10	5 meters
R2UL to R5SS	R2UL-P33 to R5SS-P6 R2UL-P34 to R5SS-P8	10 meters
R2UL to R6SS	R2UL-P35 to R6SS-P6 R2UL-P36 to R6SS-P8	10 meters
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22	5 meters
R2LL to R2SS	R2LL-P3 to R2SS-P20 R2LL-P4 to R2SS-P22	3 meters
R2LL to R3SS	R2LL-P5 to R3SS-P20 R2LL-P6 to R3SS-P22	5 meters
R2LL to R4SS	R2LL-P7 to R4SS-P20 R2LL-P30 to R4SS-P22 R2LL-P31 to R4SS-P24	5 meters

Table 4-16 (Cont.) Leaf Switch Connections for the Second Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R2LL to R5SS	R2LL-P32 to R5SS-P20 R2LL-P33 to R5SS-P22 R2LL-P34 to R5SS-P24	10 meters
R2LL to R6SS	R2LL-P35 to R6SS-P20 R2LL-P36 to R6SS-P22	10 meters

Table 4-17 Leaf Switch Connections for the Third Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R3UL to R1SS	R3UL-P1 to R1SS-P17 R3UL-P2 to R1SS-P18	5 meters
R3UL to R2SS	R3UL-P3 to R2SS-P17 R3UL-P4 to R2SS-P18	5 meters
R3UL to R3SS	R3UL-P5 to R3SS-P17 R3UL-P6 to R3SS-P18	3 meters
R3UL to R4SS	R3UL-P7 to R4SS-P17 R3UL-P30 to R4SS-P18	5 meters
R3UL to R5SS	R3UL-P31 to R5SS-P16 R3UL-P32 to R5SS-P17 R3UL-P33 to R5SS-P18	5 meters
R3UL to R6SS	R3UL-P34 to R6SS-P16 R3UL-P35 to R6SS-P17 R3UL-P36 to R6SS-P18	10 meters
R3LL to R1SS	R3LL-P1 to R1SS-P30 R3LL-P2 to R1SS-P31 R3LL-P3 to R1SS-P32	5 meters
R3LL to R2SS	R3LL-P4 to R2SS-P31 R3LL-P5 to R2SS-P32	5 meters
R3LL to R3SS	R3LL-P6 to R3SS-P31 R3LL-P7 to R3SS-P32	3 meters
R3LL to R4SS	R3LL-P30 to R4SS-P31 R3LL-P31 to R4SS-P32	5 meters
R3LL to R5SS	R3LL-P32 to R5SS-P31 R3LL-P33 to R5SS-P32	5 meters
R3LL to R6SS	R3LL-P34 to R6SS-P30 R3LL-P35 to R6SS-P31 R3LL-P36 to R6SS-P32	10 meters

Table 4-18 Leaf Switch Connections for the Fourth Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R4UL to R1SS	R4UL-P1 to R1SS-P13 R4UL-P2 to R1SS-P14 R4UL-P3 to R1SS-P15	10 meters
R4UL to R2SS	R4UL-P4 to R2SS-P12 R4UL-P5 to R2SS-P13 R4UL-P6 to R2SS-P14	5 meters
R4UL to R3SS	R4UL-P7 to R3SS-P13 R4UL-P30 to R3SS-P14	5 meters
R4UL to R4SS	R4UL-P31 to R4SS-P13 R4UL-P32 to R4SS-P14	3 meters
R4UL to R5SS	R4UL-P33 to R5SS-P13 R4UL-P34 to R5SS-P14	5 meters
R4UL to R6SS	R4UL-P35 to R6SS-P13 R4UL-P36 to R6SS-P14	5 meters
R4LL to R1SS	R4LL-P1 to R1SS-P27 R4LL-P2 to R1SS-P28	10 meters
R4LL to R2SS	R4LL-P3 to R2SS-P26 R4LL-P4 to R2SS-P27 R4LL-P5 to R2SS-P28	5 meters
R4LL to R3SS	R4LL-P6 to R3SS-P26 R4LL-P7 to R3SS-P27 R4LL-P30 to R3SS-P28	5 meters
R4LL to R4SS	R4LL-P31 to R4SS-P27 R4LL-P32 to R4SS-P28	3 meters
R4LL to R5SS	R4LL-P33 to R5SS-P25 R4LL-P34 to R5SS-P27	5 meters
R4LL to R6SS	R4LL-P35 to R6SS-P27 R4LL-P36 to R6SS-P28	5 meters

Table 4-19 Leaf Switch Connections for the Fifth Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R5UL to R1SS	R5UL-P1 to R1SS-P11 R5UL-P2 to R1SS-P12	10 meters
R5UL to R2SS	R5UL-P3 to R2SS-P10 R5UL-P4 to R2SS-P11	10 meters
R5UL to R3SS	R5UL-P5 to R3SS-P9 R5UL-P6 to R3SS-P11 R5UL-P7 to R3SS-P12	5 meters

Table 4-19 (Cont.) Leaf Switch Connections for the Fifth Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R5UL to R4SS	R5UL-P30 to R4SS-P9 R5UL-P31 to R4SS-P11 R5UL-P32 to R4SS-P12	5 meters
R5UL to R5SS	R5UL-P33 to R5SS-P11 R5UL-P34 to R5SS-P12	3 meters
R5UL to R6SS	R5UL-P35 to R6SS-P11 R5UL-P36 to R6SS-P12	5 meters
R5LL to R1SS	R5LL-P1 to R1SS-P25 R5LL-P2 to R1SS-P26	10 meters
R5LL to R2SS	R5LL-P3 to R2SS-P25 R5LL-P4 to R2SS-P29	10 meters
R5LL to R3SS	R5LL-P5 to R3SS-P24 R5LL-P6 to R3SS-P25	5 meters
R5LL to R4SS	R5LL-P7 to R4SS-P23 R5LL-P30 to R4SS-P25 R5LL-P31 to R4SS-P26	5 meters
R5LL to R5SS	R5LL-P32 to R5SS-P26 R5LL-P33 to R5SS-P28 R5LL-P34 to R5SS-P30	3 meters
R5LL to R6SS	R5LL-P35 to R6SS-P25 R5LL-P36 to R6SS-P26	5 meters

Table 4-20 Leaf Switch Connections for the Sixth Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R6UL to R1SS	R6UL-P1 to R1SS-P10 R6UL-P2 to R1SS-P16	10 meters
R6UL to R2SS	R6UL-P3 to R2SS-P15 R6UL-P4 to R2SS-P16	10 meters
R6UL to R3SS	R6UL-P5 to R3SS-P15 R6UL-P6 to R3SS-P16	10 meters
R6UL to R4SS	R6UL-P7 to R4SS-P15 R6UL-P30 to R4SS-P16	5 meters
R6UL to R5SS	R6UL-P31 to R5SS-P9 R6UL-P32 to R5SS-P10 R6UL-P33 to R5SS-P15	5 meters
R6UL to R6SS	R6UL-P34 to R6SS-P9 R6UL-P35 to R6SS-P10 R6UL-P36 to R6SS-P15	3 meters

Table 4-20 (Cont.) Leaf Switch Connections for the Sixth Rack in a Six-Rack System

Switch Pair	Connections	Cable Length
R6LL to R1SS	R6LL-P1 to R1SS-P23 R6LL-P2 to R1SS-P24 R6LL-P3 to R1SS-P29	10 meters
R6LL to R2SS	R6LL-P4 to R2SS-P24 R6LL-P5 to R2SS-P30	10 meters
R6LL to R3SS	R6LL-P6 to R3SS-P29 R6LL-P7 to R3SS-P30	10 meters
R6LL to R4SS	R6LL-P30 to R4SS-P29 R6LL-P31 to R4SS-P30	5 meters
R6LL to R5SS	R6LL-P32 to R5SS-P23 R6LL-P33 to R5SS-P29	5 meters
R6LL to R6SS	R6LL-P34 to R6SS-P23 R6LL-P35 to R6SS-P24 R6LL-P36 to R6SS-P29	3 meters

4.8 Seven-Rack Cabling for X9M Racks

This section provides the cabling details to connect seven (7) X9M racks together, all of which use RoCE Network Fabric.

 **Note:**

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following diagrams show the cable connections for the spine switches when cabling seven racks together:

Rack 1 Spine (R1SS)

1	3	5 R1 UL P1	7 R1 UL P2	9 R7 UL P1	11 R5 UL P1	13 R4 UL P1	15 R7 UL P2	17 R3 UL P1	19 R1 LL P1	21 R1 LL P2	23 R6 LL P1	25 R5 LL P1	27 R4 LL P1	29 R7 LL P1	31 R3 LL P1	33	35
2	4	6 R2 UL P1	8 R2 UL P2	10 R6 UL P1	12 R5 UL P2	14 R4 UL P2	16 R6 UL P2	18 R3 UL P2	20 R2 LL P1	22 R2 LL P2	24 R6 LL P2	26 R5 LL P2	28 R4 LL P2	30 R7 LL P2	32 R3 LL P2	34	36

Rack 2 Spine (R2SS)

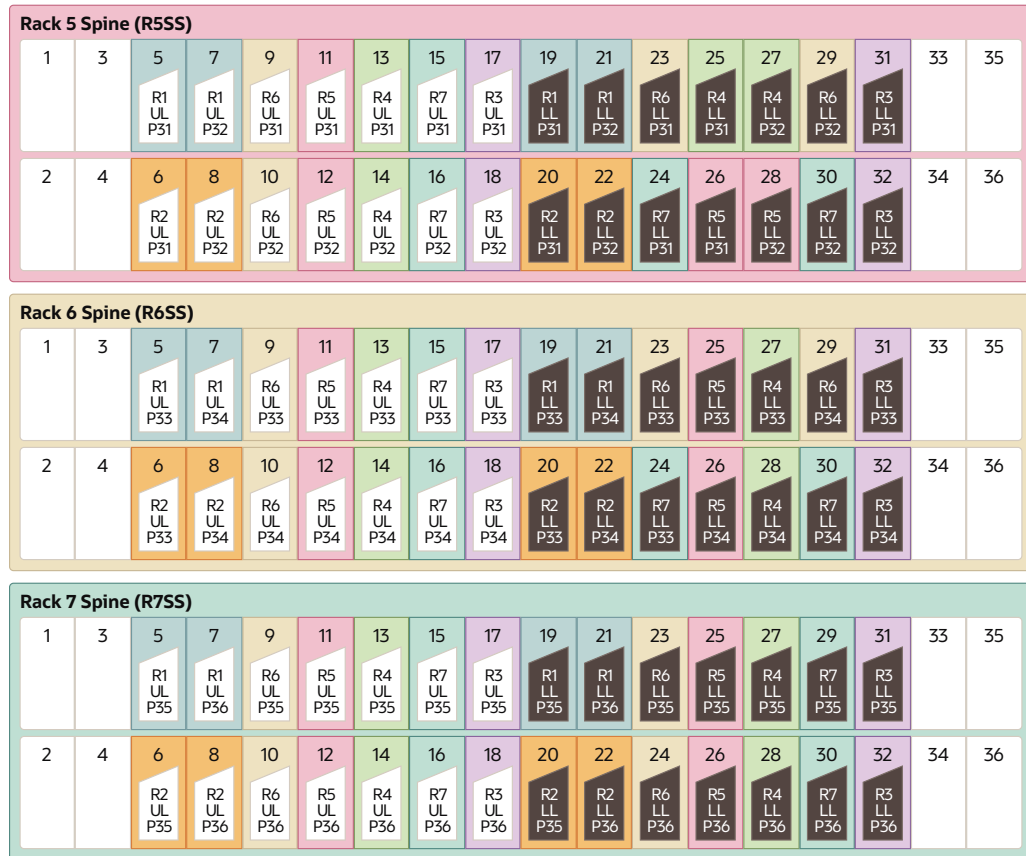
1	3	5 R1 UL P3	7 R1 UL P4	9 R7 UL P3	11 R5 UL P4	13 R4 UL P3	15 R6 UL P3	17 R3 UL P3	19 R1 LL P3	21 R1 LL P4	23 R7 LL P3	25 R5 LL P3	27 R4 LL P3	29 R5 LL P4	31 R3 LL P3	33	35
2	4	6 R2 UL P3	8 R2 UL P4	10 R5 UL P3	12 R7 UL P4	14 R4 UL P4	16 R6 UL P4	18 R3 UL P4	20 R2 LL P3	22 R2 LL P4	24 R6 LL P3	26 R7 LL P4	28 R4 LL P4	30 R6 LL P4	32 R3 LL P4	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P5	7 R1 UL P6	9 R7 UL P5	11 R5 UL P5	13 R4 UL P5	15 R6 UL P5	17 R3 UL P5	19 R1 LL P5	21 R1 LL P6	23 R7 LL P5	25 R5 LL P6	27 R4 LL P5	29 R6 LL P5	31 R3 LL P5	33	35
2	4	6 R2 UL P5	8 R2 UL P6	10 R7 UL P6	12 R5 UL P6	14 R4 UL P6	16 R6 UL P6	18 R3 UL P6	20 R2 LL P5	22 R2 LL P6	24 R5 LL P5	26 R7 LL P6	28 R4 LL P6	30 R6 LL P6	32 R3 LL P6	34	36

Rack 4 Spine (R4SS)

1	3	5 R1 UL P7	7 R1 UL P30	9 R7 UL P7	11 R5 UL P7	13 R4 UL P7	15 R6 UL P7	17 R3 UL P7	19 R1 LL P7	21 R1 LL P30	23 R7 LL P7	25 R5 LL P7	27 R4 LL P7	29 R6 LL P7	31 R3 LL P7	33	35
2	4	6 R2 UL P7	8 R2 UL P30	10 R7 UL P30	12 R5 UL P30	14 R4 UL P30	16 R6 UL P30	18 R3 UL P30	20 R2 LL P7	22 R2 LL P30	24 R7 LL P30	26 R5 LL P30	28 R4 LL P30	30 R6 LL P30	32 R3 LL P30	34	36



The following tables contain details for all of the RoCE Network Fabric cabling connections in a seven-rack system.

Table 4-21 Leaf Switch Connections for the First Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7	3 meters
R1UL to R2SS	R1UL-P3 to R2SS-P5 R1UL-P4 to R2SS-P7	5 meters
R1UL to R3SS	R1UL-P5 to R3SS-P5 R1UL-P6 to R3SS-P7	5 meters
R1UL to R4SS	R1UL-P7 to R4SS-P5 R1UL-P30 to R4SS-P7	10 meters
R1UL to R5SS	R1UL-P31 to R5SS-P5 R1UL-P32 to R5SS-P7	10 meters
R1UL to R6SS	R1UL-P33 to R6SS-P5 R1UL-P34 to R6SS-P7	10 meters
R1UL to R7SS	R1UL-P35 to R7SS-P5 R1UL-P36 to R7SS-P7	10 meters
R1LL to R1SS	R1LL-P1 to R1SS-P19 R1LL-P2 to R1SS-P21	3 meters

Table 4-21 (Cont.) Leaf Switch Connections for the First Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R1LL to R2SS	R1LL-P3 to R2SS-P19 R1LL-P4 to R2SS-P21	5 meters
R1LL to R3SS	R1LL-P5 to R3SS-P19 R1LL-P6 to R3SS-P21	5 meters
R1LL to R4SS	R1LL-P7 to R4SS-P19 R1LL-P30 to R4SS-P21	10 meters
R1LL to R5SS	R1LL-P31 to R5SS-P19 R1LL-P32 to R5SS-P21	10 meters
R1LL to R6SS	R1LL-P33 to R6SS-P19 R1LL-P34 to R6SS-P21	10 meters
R1LL to R7SS	R1LL-P35 to R7SS-P19 R1LL-P36 to R7SS-P21	10 meters

Table 4-22 Leaf Switch Connections for the Second Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6 R2UL-P2 to R1SS-P8	5 meters
R2UL to R2SS	R2UL-P3 to R2SS-P6 R2UL-P4 to R2SS-P8	3 meters
R2UL to R3SS	R2UL-P5 to R3SS-P6 R2UL-P6 to R3SS-P8	5 meters
R2UL to R4SS	R2UL-P7 to R4SS-P6 R2UL-P30 to R4SS-P8	5 meters
R2UL to R5SS	R2UL-P31 to R5SS-P6 R2UL-P32 to R5SS-P8	10 meters
R2UL to R6SS	R2UL-P33 to R6SS-P6 R2UL-P34 to R6SS-P8	10 meters
R2UL to R7SS	R2UL-P35 to R7SS-P6 R2UL-P36 to R7SS-P8	10 meters
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22	5 meters
R2LL to R2SS	R2LL-P3 to R2SS-P20 R2LL-P4 to R2SS-P22	3 meters
R2LL to R3SS	R2LL-P5 to R3SS-P20 R2LL-P6 to R3SS-P22	5 meters
R2LL to R4SS	R2LL-P7 to R4SS-P20 R2LL-P30 to R4SS-P22	5 meters

Table 4-22 (Cont.) Leaf Switch Connections for the Second Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R2LL to R5SS	R2LL-P31 to R5SS-P20 R2LL-P32 to R5SS-P22	10 meters
R2LL to R6SS	R2LL-P33 to R6SS-P20 R2LL-P34 to R6SS-P22	10 meters
R2LL to R7SS	R2LL-P35 to R7SS-P20 R2LL-P36 to R7SS-P22	10 meters

Table 4-23 Leaf Switch Connections for the Third Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R3UL to R1SS	R3UL-P1 to R1SS-P17 R3UL-P2 to R1SS-P18	5 meters
R3UL to R2SS	R3UL-P3 to R2SS-P17 R3UL-P4 to R2SS-P18	5 meters
R3UL to R3SS	R3UL-P5 to R3SS-P17 R3UL-P6 to R3SS-P18	3 meters
R3UL to R4SS	R3UL-P7 to R4SS-P17 R3UL-P30 to R4SS-P18	5 meters
R3UL to R5SS	R3UL-P31 to R5SS-P17 R3UL-P32 to R5SS-P18	5 meters
R3UL to R6SS	R3UL-P33 to R6SS-P17 R3UL-P34 to R6SS-P18	10 meters
R3UL to R7SS	R3UL-P35 to R7SS-P17 R3UL-P36 to R7SS-P18	10 meters
R3LL to R1SS	R3LL-P1 to R1SS-P31 R3LL-P2 to R1SS-P32	5 meters
R3LL to R2SS	R3LL-P3 to R2SS-P31 R3LL-P4 to R2SS-P32	5 meters
R3LL to R3SS	R3LL-P5 to R3SS-P31 R3LL-P6 to R3SS-P32	3 meters
R3LL to R4SS	R3LL-P7 to R4SS-P31 R3LL-P30 to R4SS-P32	5 meters
R3LL to R5SS	R3LL-P31 to R5SS-P31 R3LL-P32 to R5SS-P32	5 meters
R3LL to R6SS	R3LL-P33 to R6SS-P31 R3LL-P34 to R6SS-P32	10 meters
R3LL to R7SS	R3LL-P35 to R7SS-P31 R3LL-P36 to R7SS-P32	10 meters

Table 4-24 Leaf Switch Connections for the Fourth Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R4UL to R1SS	R4UL-P1 to R1SS-P13 R4UL-P2 to R1SS-P14	10 meters
R4UL to R2SS	R4UL-P3 to R2SS-P13 R4UL-P4 to R2SS-P14	5 meters
R4UL to R3SS	R4UL-P5 to R3SS-P13 R4UL-P6 to R3SS-P14	5 meters
R4UL to R4SS	R4UL-P7 to R4SS-P13 R4UL-P30 to R4SS-P14	3 meters
R4UL to R5SS	R4UL-P31 to R5SS-P13 R4UL-P32 to R5SS-P14	5 meters
R4UL to R6SS	R4UL-P33 to R6SS-P13 R4UL-P34 to R6SS-P14	5 meters
R4UL to R7SS	R4UL-P35 to R7SS-P13 R4UL-P36 to R7SS-P14	10 meters
R4LL to R1SS	R4LL-P1 to R1SS-P27 R4LL-P2 to R1SS-P28	10 meters
R4LL to R2SS	R4LL-P3 to R2SS-P27 R4LL-P4 to R2SS-P28	5 meters
R4LL to R3SS	R4LL-P5 to R3SS-P27 R4LL-P6 to R3SS-P28	5 meters
R4LL to R4SS	R4LL-P7 to R4SS-P27 R4LL-P30 to R4SS-P28	3 meters
R4LL to R5SS	R4LL-P31 to R5SS-P25 R4LL-P32 to R5SS-P27	5 meters
R4LL to R6SS	R4LL-P33 to R6SS-P27 R4LL-P34 to R6SS-P28	5 meters
R4LL to R7SS	R4LL-P35 to R7SS-P27 R4LL-P36 to R7SS-P28	10 meters

Table 4-25 Leaf Switch Connections for the Fifth Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R5UL to R1SS	R5UL-P1 to R1SS-P11 R5UL-P2 to R1SS-P12	10 meters
R5UL to R2SS	R5UL-P3 to R2SS-P10 R5UL-P4 to R2SS-P11	10 meters
R5UL to R3SS	R5UL-P5 to R3SS-P11 R5UL-P6 to R3SS-P12	5 meters
R5UL to R4SS	R5UL-P7 to R4SS-P11 R5UL-P30 to R4SS-P12	5 meters

Table 4-25 (Cont.) Leaf Switch Connections for the Fifth Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R5UL to R5SS	R5UL-P31 to R5SS-P11 R5UL-P32 to R5SS-P12	3 meters
R5UL to R6SS	R5UL-P33 to R6SS-P11 R5UL-P34 to R6SS-P12	5 meters
R5UL to R7SS	R5UL-P35 to R7SS-P11 R5UL-P36 to R7SS-P12	5 meters
R5LL to R1SS	R5LL-P1 to R1SS-P25 R5LL-P2 to R1SS-P26	10 meters
R5LL to R2SS	R5LL-P3 to R2SS-P25 R5LL-P4 to R2SS-P29	10 meters
R5LL to R3SS	R5LL-P5 to R3SS-P24 R5LL-P6 to R3SS-P25	5 meters
R5LL to R4SS	R5LL-P7 to R4SS-P25 R5LL-P30 to R4SS-P26	5 meters
R5LL to R5SS	R5LL-P31 to R5SS-P26 R5LL-P32 to R5SS-P28	3 meters
R5LL to R6SS	R5LL-P33 to R6SS-P25 R5LL-P34 to R6SS-P26	5 meters
R5LL to R7SS	R5LL-P35 to R7SS-P25 R5LL-P36 to R7SS-P26	5 meters

Table 4-26 Leaf Switch Connections for the Sixth Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R6UL to R1SS	R6UL-P1 to R1SS-P10 R6UL-P2 to R1SS-P16	10 meters
R6UL to R2SS	R6UL-P3 to R2SS-P15 R6UL-P4 to R2SS-P16	10 meters
R6UL to R3SS	R6UL-P5 to R3SS-P15 R6UL-P6 to R3SS-P16	10 meters
R6UL to R4SS	R6UL-P7 to R4SS-P15 R6UL-P30 to R4SS-P16	5 meters
R6UL to R5SS	R6UL-P31 to R5SS-P9 R6UL-P32 to R5SS-P10	5 meters
R6UL to R6SS	R6UL-P33 to R6SS-P9 R6UL-P34 to R6SS-P10	3 meters
R6UL to R7SS	R6UL-P35 to R7SS-P9 R6UL-P36 to R7SS-P10	5 meters

Table 4-26 (Cont.) Leaf Switch Connections for the Sixth Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R6LL to R1SS	R6LL-P1 to R1SS-P23 R6LL-P2 to R1SS-P24	10 meters
R6LL to R2SS	R6LL-P3 to R2SS-P24 R6LL-P4 to R2SS-P30	10 meters
R6LL to R3SS	R6LL-P5 to R3SS-P29 R6LL-P6 to R3SS-P30	10 meters
R6LL to R4SS	R6LL-P7 to R4SS-P29 R6LL-P30 to R4SS-P30	5 meters
R6LL to R5SS	R6LL-P31 to R5SS-P23 R6LL-P32 to R5SS-P29	5 meters
R6LL to R6SS	R6LL-P33 to R6SS-P23 R6LL-P34 to R6SS-P29	3 meters
R6LL to R7SS	R6LL-P35 to R7SS-P23 R6LL-P36 to R7SS-P24	5 meters

Table 4-27 Leaf Switch Connections for the Seventh Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R7UL to R1SS	R7UL-P1 to R1SS-P9 R7UL-P2 to R1SS-P15	10 meters
R7UL to R2SS	R7UL-P3 to R2SS-P9 R7UL-P4 to R2SS-P12	10 meters
R7UL to R3SS	R7UL-P5 to R3SS-P9 R7UL-P6 to R3SS-P10	10 meters
R7UL to R4SS	R7UL-P7 to R4SS-P9 R7UL-P30 to R4SS-P10	10 meters
R7UL to R5SS	R7UL-P31 to R5SS-P15 R7UL-P32 to R5SS-P16	5 meters
R7UL to R6SS	R7UL-P33 to R6SS-P15 R7UL-P34 to R6SS-P16	5 meters
R7UL to R7SS	R7UL-P35 to R7SS-P15 R7UL-P36 to R7SS-P16	3 meters
R7LL to R1SS	R7LL-P1 to R1SS-P29 R7LL-P2 to R1SS-P30	10 meters
R7LL to R2SS	R7LL-P3 to R2SS-P23 R7LL-P4 to R2SS-P26	10 meters
R7LL to R3SS	R7LL-P5 to R3SS-P23 R7LL-P6 to R3SS-P26	10 meters

Table 4-27 (Cont.) Leaf Switch Connections for the Seventh Rack in a Seven-Rack System

Switch Pair	Connections	Cable Length
R7LL to R4SS	R7LL-P7 to R4SS-P23 R7LL-P30 to R4SS-P24	10 meters
R7LL to R5SS	R7LL-P31 to R5SS-P24 R7LL-P32 to R5SS-P30	5 meters
R7LL to R6SS	R7LL-P33 to R6SS-P24 R7LL-P34 to R6SS-P30	5 meters
R7LL to R7SS	R7LL-P35 to R7SS-P29 R7LL-P36 to R7SS-P30	3 meters

4.9 Eight-Rack Cabling for X9M Racks

This section provides the cabling details to connect eight (8) X9M racks together, all of which use RoCE Network Fabric.

 **Note:**

- The following conventions are used in the cabling notation for connecting multiple racks together:
 - The abbreviation for the first rack is R1, the second rack is R2, and so on.
 - LL identifies a lower leaf switch and UL identifies an upper leaf switch.
 - SS identifies the spine switch, which is located in U1 on all racks.
 - A specific switch is identified by combining abbreviations. For example, R1LL identifies the lower leaf switch (LL) on the first rack (R1).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X9M-2 or Storage Expansion Rack X9M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X9M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X9M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must be deleted, leaving only 13 inter-switch links on each leaf switch. These changes only apply to X9M-8 systems with three database servers and 11 storage servers.

The following diagrams show the cable connections for the spine switches when cabling eight racks together:

Rack 1 Spine (R1SS)

1	3	5 R1 UL P1	7 R1 UL P2	9 R7 UL P1	11 R5 UL P1	13 R4 UL P1	15 R7 UL P2	17 R3 UL P1	19 R1 LL P1	21 R8 LL P1	23 R6 LL P1	25 R8 LL P2	27 R4 LL P1	29 R7 LL P1	31 R3 LL P1	33	35
2	4	6 R2 UL P1	8 R8 UL P1	10 R6 UL P1	12 R5 UL P2	14 R4 UL P2	16 R8 UL P2	18 R3 UL P2	20 R2 LL P1	22 R2 LL P2	24 R6 LL P2	26 R5 LL P1	28 R4 LL P2	30 R7 LL P2	32 R3 LL P2	34	36

Rack 2 Spine (R2SS)

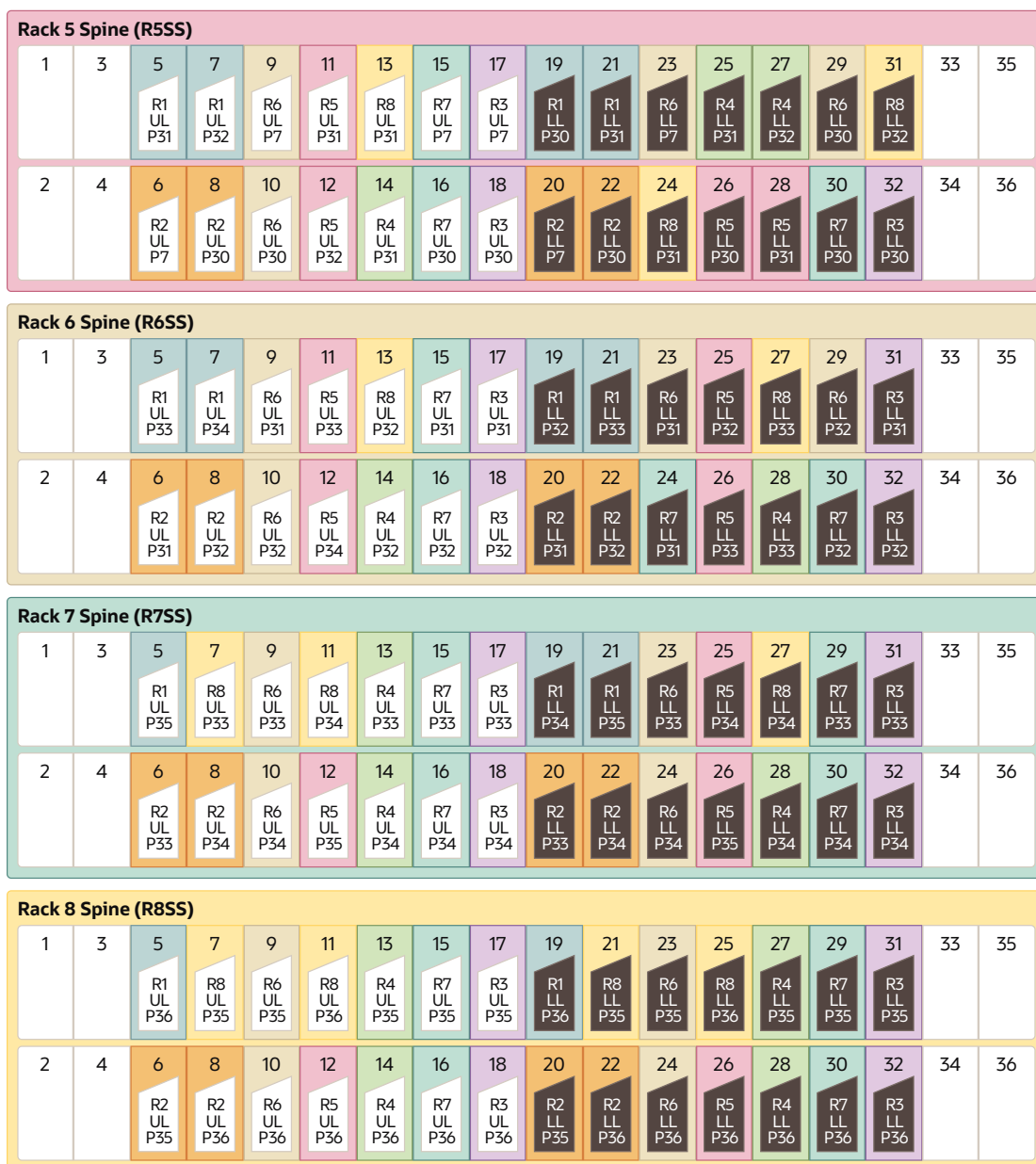
1	3	5 R1 LL P3	7 R1 UL P4	9 R7 UL P3	11 R5 UL P4	13 R4 UL P3	15 R6 UL P2	17 R3 UL P3	19 R1 LL P2	21 R1 LL P3	23 R7 LL P3	25 R5 LL P2	27 R4 LL P3	29 R5 LL P3	31 R3 LL P3	33	35
2	4	6 R2 UL P2	8 R8 UL P3	10 R5 UL P3	12 R7 UL P4	14 R4 UL P4	16 R8 UL P4	18 R3 UL P4	20 R2 LL P3	22 R8 LL P3	24 R6 LL P3	26 R7 LL P4	28 R4 LL P4	30 R8 LL P4	32 R3 LL P4	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P5	7 R1 UL P6	9 R7 UL P5	11 R5 UL P5	13 R4 UL P5	15 R6 UL P3	17 R8 UL P6	19 R1 LL P4	21 R1 LL P5	23 R7 LL P5	25 R5 LL P5	27 R4 LL P5	29 R8 LL P6	31 R3 LL P5	33	35
2	4	6 R2 UL P3	8 R2 UL P4	10 R8 UL P5	12 R5 UL P6	14 R4 UL P6	16 R6 UL P4	18 R3 UL P5	20 R2 LL P4	22 R8 LL P5	24 R5 LL P4	26 R7 LL P6	28 R4 LL P6	30 R6 LL P4	32 R3 LL P6	34	36

Rack 4 Spine (R4SS)

1	3	5 R1 UL P7	7 R1 UL P30	9 R7 UL P6	11 R5 UL P7	13 R4 UL P7	15 R6 UL P5	17 R8 UL P30	19 R1 LL P6	21 R1 LL P7	23 R8 LL P7	25 R5 LL P6	27 R4 LL P7	29 R6 LL P5	31 R8 LL P30	33	35
2	4	6 R2 UL P5	8 R2 UL P6	10 R8 UL P7	12 R5 UL P30	14 R4 UL P30	16 R4 UL P6	18 R3 UL P6	20 R2 LL P5	22 R2 LL P6	24 R7 LL P7	26 R5 LL P7	28 R4 LL P30	30 R6 LL P6	32 R3 LL P7	34	36



The following tables contain details for all of the RoCE Network Fabric cabling connections in an eight-rack system.

Table 4-28 Leaf Switch Connections for the First Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R1UL to R1SS	R1UL-P1 to R1SS-P5 R1UL-P2 to R1SS-P7	3 meters
R1UL to R2SS	R1UL-P3 to R2SS-P5 R1UL-P4 to R2SS-P7	5 meters
R1UL to R3SS	R1UL-P5 to R3SS-P5 R1UL-P6 to R3SS-P7	5 meters

Table 4-28 (Cont.) Leaf Switch Connections for the First Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R1UL to R4SS	R1UL-P7 to R4SS-P5 R1UL-P30 to R4SS-P7	10 meters
R1UL to R5SS	R1UL-P31 to R5SS-P5 R1UL-P32 to R5SS-P7	10 meters
R1UL to R6SS	R1UL-P33 to R6SS-P5 R1UL-P34 to R6SS-P7	10 meters
R1UL to R7SS	R1UL-P35 to R7SS-P5	10 meters
R1UL to R8SS	R1UL-P36 to R8SS-P5	10 meters
R1LL to R1SS	R1LL-P1 to R1SS-P19	3 meters
R1LL to R2SS	R1LL-P2 to R2SS-P19 R1LL-P3 to R2SS-P21	5 meters
R1LL to R3SS	R1LL-P4 to R3SS-P19 R1LL-P5 to R3SS-P21	5 meters
R1LL to R4SS	R1LL-P6 to R4SS-P19 R1LL-P7 to R4SS-P21	10 meters
R1LL to R5SS	R1LL-P30 to R5SS-P19 R1LL-P31 to R5SS-P21	10 meters
R1LL to R6SS	R1LL-P32 to R6SS-P19 R1LL-P33 to R6SS-P21	10 meters
R1LL to R7SS	R1LL-P34 to R7SS-P19 R1LL-P35 to R7SS-P21	10 meters
R1LL to R8SS	R1LL-P36 to R8SS-P19	10 meters

Table 4-29 Leaf Switch Connections for the Second Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R2UL to R1SS	R2UL-P1 to R1SS-P6	5 meters
R2UL to R2SS	R2UL-P2 to R2SS-P6	3 meters
R2UL to R3SS	R2UL-P3 to R3SS-P6 R2UL-P4 to R3SS-P8	5 meters
R2UL to R4SS	R2UL-P5 to R4SS-P6 R2UL-P6 to R4SS-P8	5 meters
R2UL to R5SS	R2UL-P7 to R5SS-P6 R2UL-P30 to R5SS-P8	10 meters
R2UL to R6SS	R2UL-P31 to R6SS-P6 R2UL-P32 to R6SS-P8	10 meters
R2UL to R7SS	R2UL-P33 to R7SS-P6 R2UL-P34 to R7SS-P8	10 meters

Table 4-29 (Cont.) Leaf Switch Connections for the Second Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R2UL to R8SS	R2UL-P35 to R8SS-P6 R2UL-P36 to R8SS-P8	10 meters
R2LL to R1SS	R2LL-P1 to R1SS-P20 R2LL-P2 to R1SS-P22	5 meters
R2LL to R2SS	R2LL-P3 to R2SS-P20	3 meters
R2LL to R3SS	R2LL-P4 to R3SS-P20	5 meters
R2LL to R4SS	R2LL-P5 to R4SS-P20 R2LL-P6 to R4SS-P22	5 meters
R2LL to R5SS	R2LL-P7 to R5SS-P20 R2LL-P30 to R5SS-P22	10 meters
R2LL to R6SS	R2LL-P31 to R6SS-P20 R2LL-P32 to R6SS-P22	10 meters
R2LL to R7SS	R2LL-P33 to R7SS-P20 R2LL-P34 to R7SS-P22	10 meters
R2LL to R8SS	R2LL-P35 to R8SS-P20 R2LL-P36 to R8SS-P22	10 meters

Table 4-30 Leaf Switch Connections for the Third Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R3UL to R1SS	R3UL-P1 to R1SS-P17 R3UL-P2 to R1SS-P18	5 meters
R3UL to R2SS	R3UL-P3 to R2SS-P17 R3UL-P4 to R2SS-P18	5 meters
R3UL to R3SS	R3UL-P5 to R3SS-P18	3 meters
R3UL to R4SS	R3UL-P6 to R4SS-P18	5 meters
R3UL to R5SS	R3UL-P7 to R5SS-P17 R3UL-P30 to R5SS-P18	5 meters
R3UL to R6SS	R3UL-P31 to R6SS-P17 R3UL-P32 to R6SS-P18	10 meters
R3UL to R7SS	R3UL-P33 to R7SS-P17 R3UL-P34 to R7SS-P18	10 meters
R3UL to R8SS	R3UL-P35 to R8SS-P17 R3UL-P36 to R8SS-P18	10 meters
R3LL to R1SS	R3LL-P1 to R1SS-P31 R3LL-P2 to R1SS-P32	5 meters
R3LL to R2SS	R3LL-P3 to R2SS-P31 R3LL-P4 to R2SS-P32	5 meters

Table 4-30 (Cont.) Leaf Switch Connections for the Third Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R3LL to R3SS	R3LL-P5 to R3SS-P31 R3LL-P6 to R3SS-P32	3 meters
R3LL to R4SS	R3LL-P7 to R4SS-P32	5 meters
R3LL to R5SS	R3LL-P30 to R5SS-P32	5 meters
R3LL to R6SS	R3LL-P31 to R6SS-P31 R3LL-P32 to R6SS-P32	10 meters
R3LL to R7SS	R3LL-P33 to R7SS-P31 R3LL-P34 to R7SS-P32	10 meters
R3LL to R8SS	R3LL-P35 to R8SS-P31 R3LL-P36 to R8SS-P32	10 meters

Table 4-31 Leaf Switch Connections for the Fourth Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R4UL to R1SS	R4UL-P1 to R1SS-P13 R4UL-P2 to R1SS-P14	10 meters
R4UL to R2SS	R4UL-P3 to R2SS-P13 R4UL-P4 to R2SS-P14	5 meters
R4UL to R3SS	R4UL-P5 to R3SS-P13 R4UL-P6 to R3SS-P14	5 meters
R4UL to R4SS	R4UL-P7 to R4SS-P13 R4UL-P30 to R4SS-P14	3 meters
R4UL to R5SS	R4UL-P31 to R5SS-P14	5 meters
R4UL to R6SS	R4UL-P32 to R6SS-P14	5 meters
R4UL to R7SS	R4UL-P33 to R7SS-P13 R4UL-P34 to R7SS-P14	10 meters
R4UL to R8SS	R4UL-P35 to R8SS-P13 R4UL-P36 to R8SS-P14	10 meters
R4LL to R1SS	R4LL-P1 to R1SS-P27 R4LL-P2 to R1SS-P28	10 meters
R4LL to R2SS	R4LL-P3 to R2SS-P27 R4LL-P4 to R2SS-P28	5 meters
R4LL to R3SS	R4LL-P5 to R3SS-P27 R4LL-P6 to R3SS-P28	5 meters
R4LL to R4SS	R4LL-P7 to R4SS-P27 R4LL-P30 to R4SS-P28	3 meters
R4LL to R5SS	R4LL-P31 to R5SS-P25 R4LL-P32 to R5SS-P27	5 meters

Table 4-31 (Cont.) Leaf Switch Connections for the Fourth Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R4LL to R6SS	R4LL-P33 to R6SS-P28	5 meters
R4LL to R7SS	R4LL-P34 to R7SS-P28	10 meters
R4LL to R8SS	R4LL-P35 to R8SS-P27 R4LL-P36 to R8SS-P28	10 meters

Table 4-32 Leaf Switch Connections for the Fifth Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R5UL to R1SS	R5UL-P1 to R1SS-P11 R5UL-P2 to R1SS-P12	10 meters
R5UL to R2SS	R5UL-P3 to R2SS-P10 R5UL-P4 to R2SS-P11	10 meters
R5UL to R3SS	R5UL-P5 to R3SS-P11 R5UL-P6 to R3SS-P12	5 meters
R5UL to R4SS	R5UL-P7 to R4SS-P11 R5UL-P30 to R4SS-P12	5 meters
R5UL to R5SS	R5UL-P31 to R5SS-P11 R5UL-P32 to R5SS-P12	3 meters
R5UL to R6SS	R5UL-P33 to R6SS-P11 R5UL-P34 to R6SS-P12	5 meters
R5UL to R7SS	R5UL-P35 to R7SS-P12	5 meters
R5UL to R8SS	R5UL-P36 to R8SS-P12	10 meters
R5LL to R1SS	R5LL-P1 to R1SS-P26	10 meters
R5LL to R2SS	R5LL-P2 to R2SS-P25 R5LL-P3 to R2SS-P29	10 meters
R5LL to R3SS	R5LL-P4 to R3SS-P24 R5LL-P5 to R3SS-P25	5 meters
R5LL to R4SS	R5LL-P6 to R4SS-P25 R5LL-P7 to R4SS-P26	5 meters
R5LL to R5SS	R5LL-P30 to R5SS-P26 R5LL-P31 to R5SS-P28	3 meters
R5LL to R6SS	R5LL-P32 to R6SS-P25 R5LL-P33 to R6SS-P26	5 meters
R5LL to R7SS	R5LL-P34 to R7SS-P25 R5LL-P35 to R7SS-P26	5 meters
R5LL to R8SS	R5LL-P36 to R8SS-P26	10 meters

Table 4-33 Leaf Switch Connections for the Sixth Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R6UL to R1SS	R6UL-P1 to R1SS-P10	10 meters
R6UL to R2SS	R6UL-P2 to R2SS-P16	10 meters
R6UL to R3SS	R6UL-P3 to R3SS-P15 R6UL-P4 to R3SS-P16	10 meters
R6UL to R4SS	R6UL-P5 to R4SS-P15 R6UL-P6 to R4SS-P16	5 meters
R6UL to R5SS	R6UL-P7 to R5SS-P9 R6UL-P30 to R5SS-P10	5 meters
R6UL to R6SS	R6UL-P31 to R6SS-P9 R6UL-P32 to R6SS-P10	3 meters
R6UL to R7SS	R6UL-P33 to R7SS-P9 R6UL-P34 to R7SS-P10	5 meters
R6UL to R8SS	R6UL-P35 to R8SS-P9 R6UL-P36 to R8SS-P10	5 meters
R6LL to R1SS	R6LL-P1 to R1SS-P23 R6LL-P2 to R1SS-P24	10 meters
R6LL to R2SS	R6LL-P3 to R2SS-P24	10 meters
R6LL to R3SS	R6LL-P4 to R3SS-P30	10 meters
R6LL to R4SS	R6LL-P5 to R4SS-P29 R6LL-P6 to R4SS-P30	5 meters
R6LL to R5SS	R6LL-P7 to R5SS-P23 R6LL-P30 to R5SS-P29	5 meters
R6LL to R6SS	R6LL-P31 to R6SS-P23 R6LL-P32 to R6SS-P29	3 meters
R6LL to R7SS	R6LL-P33 to R7SS-P23 R6LL-P34 to R7SS-P24	5 meters
R6LL to R8SS	R6LL-P35 to R8SS-P23 R6LL-P36 to R8SS-P24	5 meters

Table 4-34 Leaf Switch Connections for the Seventh Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R7UL to R1SS	R7UL-P1 to R1SS-P9 R7UL-P2 to R1SS-P15	10 meters
R7UL to R2SS	R7UL-P3 to R2SS-P9 R7UL-P4 to R2SS-P12	10 meters
R7UL to R3SS	R7UL-P5 to R3SS-P9	10 meters
R7UL to R4SS	R7UL-P6 to R4SS-P9	10 meters

Table 4-34 (Cont.) Leaf Switch Connections for the Seventh Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R7UL to R5SS	R7UL-P7 to R5SS-P15 R7UL-P30 to R5SS-P16	5 meters
R7UL to R6SS	R7UL-P31 to R6SS-P15 R7UL-P32 to R6SS-P16	5 meters
R7UL to R7SS	R7UL-P33 to R7SS-P15 R7UL-P34 to R7SS-P16	3 meters
R7UL to R8SS	R7UL-P35 to R8SS-P15 R7UL-P36 to R8SS-P16	5 meters
R7LL to R1SS	R7LL-P1 to R1SS-P29 R7LL-P2 to R1SS-P30	10 meters
R7LL to R2SS	R7LL-P3 to R2SS-P23 R7LL-P4 to R2SS-P26	10 meters
R7LL to R3SS	R7LL-P5 to R3SS-P23 R7LL-P6 to R3SS-P26	10 meters
R7LL to R4SS	R7LL-P7 to R4SS-P24	10 meters
R7LL to R5SS	R7LL-P30 to R5SS-P30	5 meters
R7LL to R6SS	R7LL-P31 to R6SS-P24 R7LL-P32 to R6SS-P30	5 meters
R7LL to R7SS	R7LL-P33 to R7SS-P29 R7LL-P34 to R7SS-P30	3 meters
R7LL to R8SS	R7LL-P35 to R8SS-P29 R7LL-P36 to R8SS-P30	5 meters

Table 4-35 Leaf Switch Connections for the Eighth Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R8UL to R1SS	R8UL-P1 to R1SS-P8 R8UL-P2 to R1SS-P16	10 meters
R8UL to R2SS	R8UL-P3 to R2SS-P8 R8UL-P4 to R2SS-P16	10 meters
R8UL to R3SS	R8UL-P5 to R3SS-P10 R8UL-P6 to R3SS-P17	10 meters
R8UL to R4SS	R8UL-P7 to R4SS-P10 R8UL-P30 to R4SS-P17	10 meters
R8UL to R5SS	R8UL-P31 to R5SS-P13	10 meters
R8UL to R6SS	R8UL-P32 to R6SS-P13	5 meters
R8UL to R7SS	R8UL-P33 to R7SS-P7 R8UL-P34 to R7SS-P11	5 meters

Table 4-35 (Cont.) Leaf Switch Connections for the Eighth Rack in a Eight-Rack System

Switch Pair	Connections	Cable Length
R8UL to R8SS	R8UL-P35 to R8SS-P7 R8UL-P36 to R8SS-P11	3 meters
R8LL to R1SS	R8LL-P1 to R1SS-P21 R8LL-P2 to R1SS-P25	10 meters
R8LL to R2SS	R8LL-P3 to R2SS-P22 R8LL-P4 to R2SS-P30	10 meters
R8LL to R3SS	R8LL-P5 to R3SS-P22 R8LL-P6 to R3SS-P29	10 meters
R8LL to R4SS	R8LL-P7 to R4SS-P23 R8LL-P30 to R4SS-P31	10 meters
R8LL to R5SS	R8LL-P31 to R5SS-P24 R8LL-P32 to R5SS-P31	10 meters
R8LL to R6SS	R8LL-P33 to R6SS-P27	5 meters
R8LL to R7SS	R8LL-P34 to R7SS-P27	5 meters
R8LL to R8SS	R8LL-P35 to R8SS-P21 R8LL-P36 to R8SS-P25	3 meters

5

Multi-Rack Cabling Tables for Oracle Exadata Rack X8M

This section contains multi-rack cabling tables for Oracle Exadata Rack X8M, which uses RoCE Network Fabric.

- [Understanding Multi-Rack Cabling for X8M Racks](#)
Up to eight racks can be cabled together without external RDMA Network Fabric switches.
- [Two-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect two X8M racks together, both of which use RoCE Network Fabric.
- [Three-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect three X8M racks together using RoCE Network Fabric.
- [Four-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect four X8M racks together, all of which use RoCE Network Fabric.
- [Five-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect five (5) X8M racks together, all of which use RoCE Network Fabric.
- [Six-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect six (6) X8M racks together, all of which use RoCE Network Fabric.
- [Seven-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect seven (7) X8M racks together, all of which use RoCE Network Fabric.
- [Eight-Rack Cabling for X8M Racks](#)
This section provides the cabling tables to connect eight (8) X8M racks together, all of which use RoCE Network Fabric.

5.1 Understanding Multi-Rack Cabling for X8M Racks

Up to eight racks can be cabled together without external RDMA Network Fabric switches.

When cabling together multiple racks, each rack must contain three RDMA Network Fabric switches.

Two of the switches, known as leaf switches, connect to all of the database servers and storage servers in the rack. The leaf switches are located as follows:

- At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2.
- At rack unit 21 (U21) and 23 (U23) in Exadata Database Machine X8M-8.

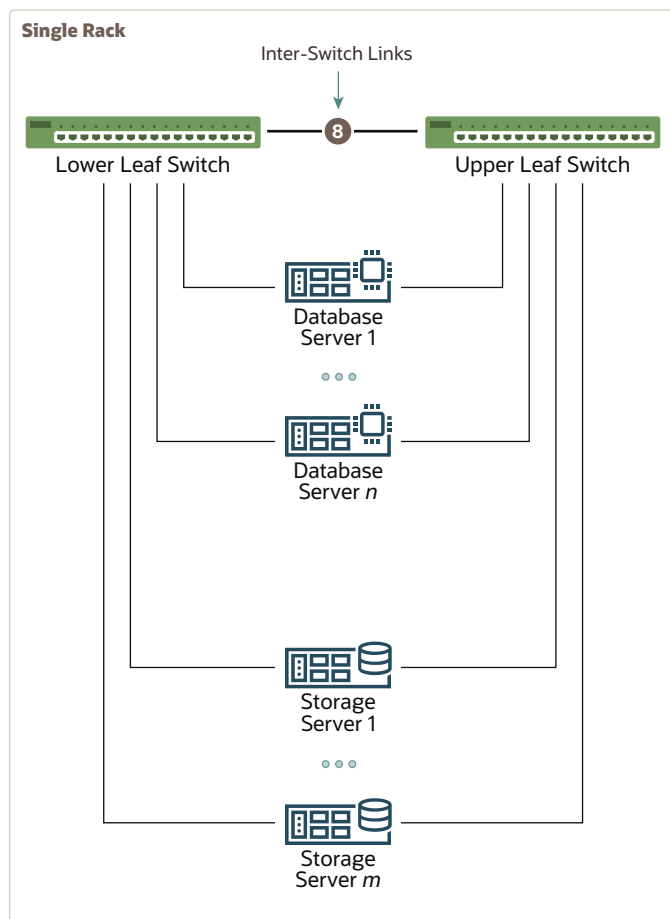
The third switch, known as the spine switch, interconnects the leaf switches in a multi-rack configuration. The spine switch is located at rack unit 1 (U1). If any rack does not include a spine switch, then order the expansion switch kit.

The procedures in this section assume the racks are adjacent to each other, standard routing in raised floor is used, and there are no obstacles in the raised floor. If these assumptions are not correct for your environment, then longer cables may be required for the connections.

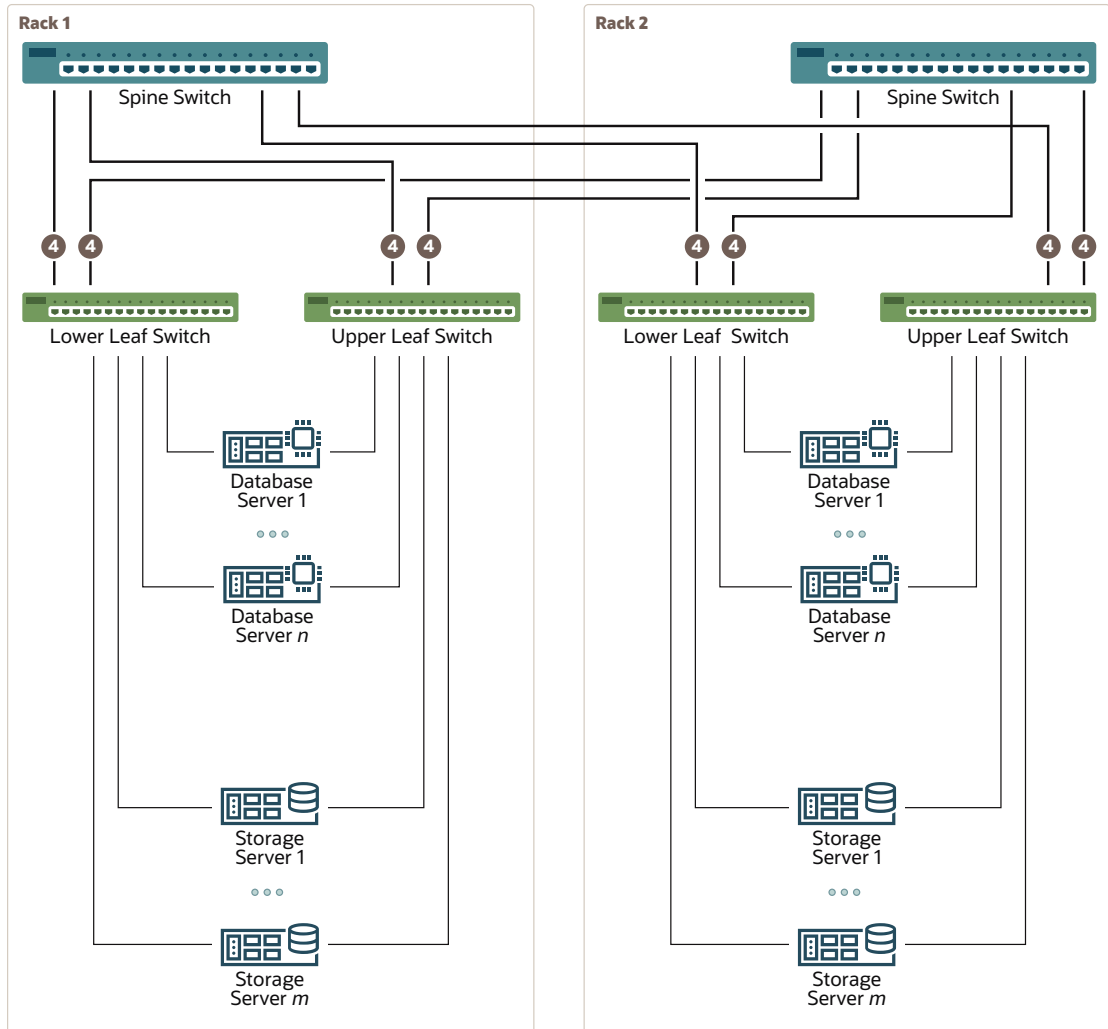
 **Note:**

By default, Oracle Exadata Database Machine X8M racks do not include spare cables or a third RoCE Network Fabric switch. To extend these racks, you must order the required cables and RoCE Network Fabric switch.

The following diagram shows the default RDMA Network Fabric architecture for a single-rack system. Each rack has two leaf switches, with eight connections between the leaf switches. The database servers and storage servers are each connected to both leaf switches. Each server contains a dual-port RDMA Network Fabric card, with port 1 connected to the lower leaf switch and port 2 connected to the upper leaf switch.



When connecting up to eight racks together, remove the eight existing inter-switch connections between each leaf switch on each rack. From each leaf switch, distribute eight connections over the spine switches in all racks. In multi-rack environments, the leaf switches inside a rack are no longer directly inter-connected, as shown in the following graphic:



As shown in the preceding graphic, each leaf switch in rack 1 has the following connections:

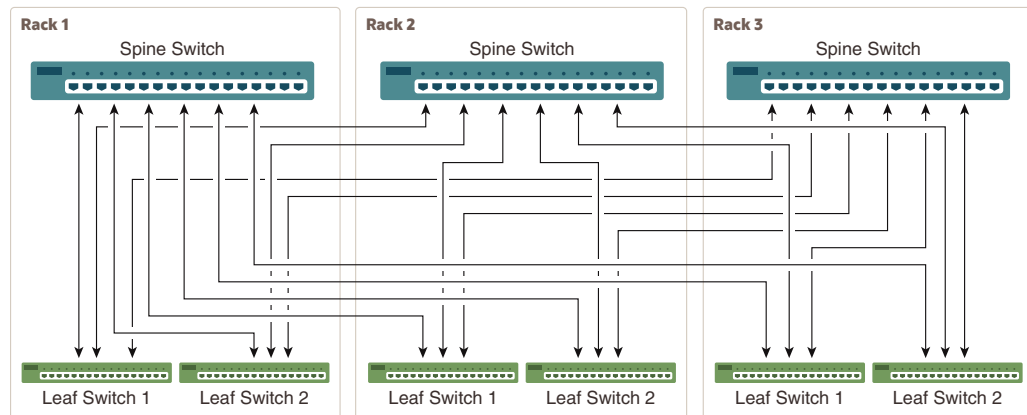
- Four connections to its internal spine switch
- Four connections to the spine switch in rack 2

The spine switch in rack 1 has the following connections:

- Four connections to each leaf switch in rack 1
- Four connections to each leaf switch in rack 2

As the number of racks increases from two to eight, the pattern continues as shown in the following graphic:

Figure 5-1 Connections Between Spine Switches and Leaf Switches for up to 8 Racks



As shown in the preceding graphic, each leaf switch has eight inter-switch connections distributed over all spine switches. Each spine switch has 16 inter-switch connections distributed over all leaf switches. The leaf switches are not directly interconnected with other leaf switches, and the spine switches are not directly interconnected with the other spine switches.

- [Preparing for Multi-Rack Cabling with X8M Racks](#)
- [Cabling Multiple Exadata Database Machine X8M Racks](#)

5.1.1 Preparing for Multi-Rack Cabling with X8M Racks

Racks can be added together to increase system capacity and performance. When cabling racks together, note the following:

- The cable lengths shown in this document assume the racks are adjacent to each other. If the racks are not adjacent, or there are obstacles in the raised floor, or if you use overhead cabling, then longer cables may be required. For optical cables, the maximum supported cable length is 100 meters. For copper cables, the maximum supported cable length is 5 meters.
- Oracle recommends that the names for the servers include the rack unit number. This helps identify the server during diagnostics.
- When completing Oracle Exadata Deployment Assistant (OEDA) for the additional rack, you are prompted for SCAN addresses. However, these SCAN addresses are not used because the SCAN address from the original rack are used. Manually remove the new SCAN addresses from the generated installation files.
- The software owner account names and group names, as well as their identifiers, must match the names and identifiers of the original rack.
- If the additional grid disks are used with existing disk groups, then ensure the grid disk sizes for the new rack are the same as the original rack.



Note:

Exadata Database Machine Eighth Racks, Quarter Racks, and Elastic Configurations are connected to other racks in the same fashion as larger racks are connected to each other. In other words, **a spine switch must exist in each rack in order to interconnect with other racks.**

Perform the following tasks before cabling racks together:

1. Determine the number of racks that will be cabled together.
2. Order the parts needed to connect the racks.

To extend Exadata Database Machine racks with RoCE Network Fabric, for each rack being added you must order extra cables, transceivers for longer cables, and a RoCE Network Fabric spine switch, if one is not already available.

When connecting four or more racks, or if you need longer cables for your environment, you must purchase additional 10 meter fiber cables with two QSFP28 SR transceivers to connect each end. The QSFP28 SR transceivers are needed for fiber cables over 5 meters in length.

The following table lists the number of racks, and cables needed to interconnect up to eight racks:

Number of Racks	3 Meter Cables Needed	5 Meter Cables Needed	10 Meter Cables Needed
2	16	16	0
3	16	32	0
4	16	40	8
5	16	46	18
6	17	49	30
7	19	51	42
8	16	52	60

3. Gather the following tools:
 - Pen
 - Note pad
 - Flashlight
 - Labels for cables and rack
 - Cable ties (Oracle recommends hook and loop straps)
4. Determine a naming method for the rack prefixes. For example, if the original rack has the prefix `dbm01`, then use the prefix `dbm02` for the second rack, the prefix `dbm03` for the third rack, and so on.
5. Verify the racks have unique host names and IP addresses. All servers interconnected in the racks must have unique names and IP addresses.

Server names and IP addresses conventions may differ in the following cases:

- **Initial installation of all Exadata Database Machine Racks:** System address assignments and host names should be complete.
 - **New Exadata Database Machine Racks are added to an existing cluster:** The new rack configuration should require unique host names and IP addresses for the new Exadata Database Machine. The IP addresses on the same subnet cannot conflict with the existing systems.
 - **Two existing Exadata Database Machine Racks are clustered together:** You can assign host names and IP addresses only if Exadata Database Machines are already assigned unique host names and IP addresses, or the entire cluster must be reconfigured. The machines must be on the same subnet and not have conflicting IP addresses.
6. Ensure the IP addresses for the new servers are in the same subnet, and do not overlap with the currently-installed servers.
 7. Check that the firmware on the original switches is at the same level as the new switches by using the `sh ver` command.

It is highly recommended, though not mandatory, to use the same firmware version on all of the switches. If the firmware is not at the same level, you can apply a firmware patch to bring the switches up to the same firmware level.

5.1.2 Cabling Multiple Exadata Database Machine X8M Racks

A spine switch must exist in the Exadata Database Machine rack, regardless of size, to multi-rack with other racks.



Note:

To extend Exadata Database Machine racks with RoCE Network Fabric, you must order cables, transceivers if needed, and a RoCE Network Fabric switch, if they are not already available.

In the following steps, the number in parentheses also indicates the number of cables required.

1. Split each leaf switch uplink bundle by the number of spine switches (or racks) in such a way that the total count adds up to 8. This split is represented in parentheses for each example.

Example 1: For two racks, take 8 uplinks from each leaf switch and split evenly by 2. Four uplinks from each leaf switch go to `rack1-spine` switch, and four uplinks from each leaf switch go to `rack2-spine` switch (4 + 4 for each leaf switch).

Example 2: For three racks, take 8 uplinks from each leaf switch and split evenly by 3. Three uplinks go to `rack1-spine` switch, three uplinks go to `rack2-spine` switch, two uplinks go to `rack3-spine` switch (3 + 3 + 2 for each leaf switch).

2. Starting from the first available port on a different spine switch, round-robin the above split scheme for each leaf switch and spine switch.

For example, for three racks:

- `rack1-leaf1` switch starts from `rack1-spine` switch for first split, `rack2-spine` switch for second split, `rack3-spine` switch for third split

- rack2-leaf1 switch starts from rack2-spine switch for first split, rack3-spine switch for second split, rack1-spine for third split
 - rack3-leaf1 switch starts from rack3-spine switch for first split, rack1-spine switch for second split, rack2-spine switch for third split
 - and so on...
3. After walking through all leaf switch uplinks in each case, you will have used all spine switch ports between port 5 and 20 inclusive.

The remaining topics in this section provide detailed cabling information for cabling up to 8 racks together.

5.2 Two-Rack Cabling for X8M Racks

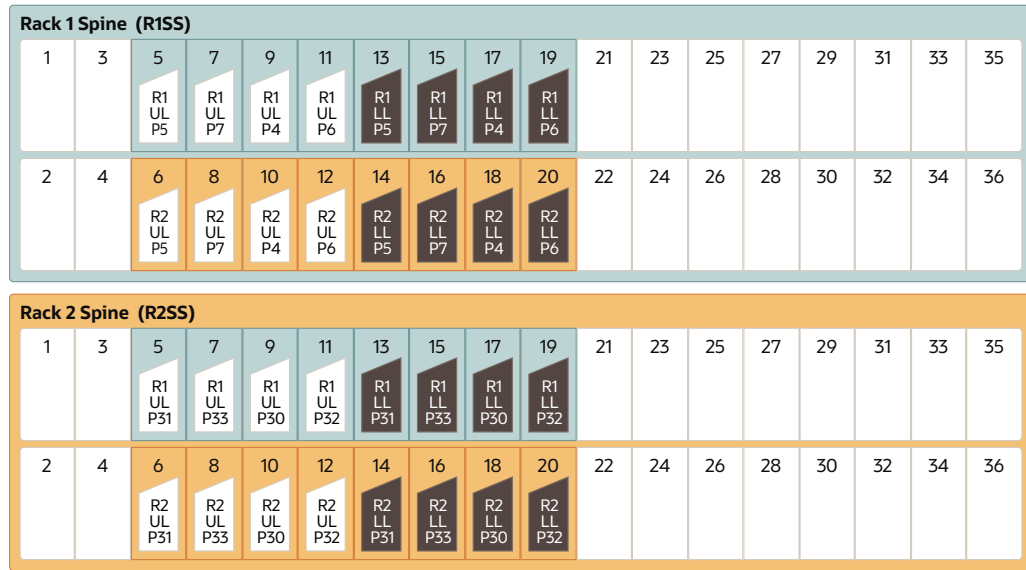
This section provides the cabling tables to connect two X8M racks together, both of which use RoCE Network Fabric.

Note:

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the two spine switches (R1 SS and R2 SS) when cabling two racks together:



The following table describes the cable connections for the first spine switch (R1 SS) when cabling two racks together:

Table 5-1 Leaf Switch Connections for the First Rack in a Two-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5 R1-UL-P7 to R1-SS-P7 R1-UL-P4 to R1-SS-P9 R1-UL-P6 to R1-SS-P11	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5 R1-UL-P33 to R2-SS-P7 R1-UL-P30 to R2-SS-P9 R1-UL-P32 to R2-SS-P11	5 meters
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13 R1-LL-P7 to R1-SS-P15 R1-LL-P4 to R1-SS-P17 R1-LL-P6 to R1-SS-P19	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13 R1-LL-P33 to R2-SS-P15 R1-LL-P30 to R2-SS-P17 R1-LL-P32 to R2-SS-P19	5 meters

The following table describes the cable connections for the second spine switch (R2 SS) when cabling two racks together:

Table 5-2 Leaf Switch Connections for the Second Rack in a Two-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6 R2-UL-P7 to R1-SS-P8 R2-UL-P4 to R1-SS-P10 R2-UL-P6 to R1-SS-P12	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6 R2-UL-P33 to R2-SS-P8 R2-UL-P30 to R2-SS-P10 R2-UL-P32 to R2-SS-P12	3 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14 R2-LL-P7 to R1-SS-P16 R2-LL-P4 to R1-SS-P18 R2-LL-P6 to R1-SS-P20	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14 R2-LL-P33 to R2-SS-P16 R2-LL-P30 to R2-SS-P18 R2-LL-P32 to R2-SS-P20	3 meters

5.3 Three-Rack Cabling for X8M Racks

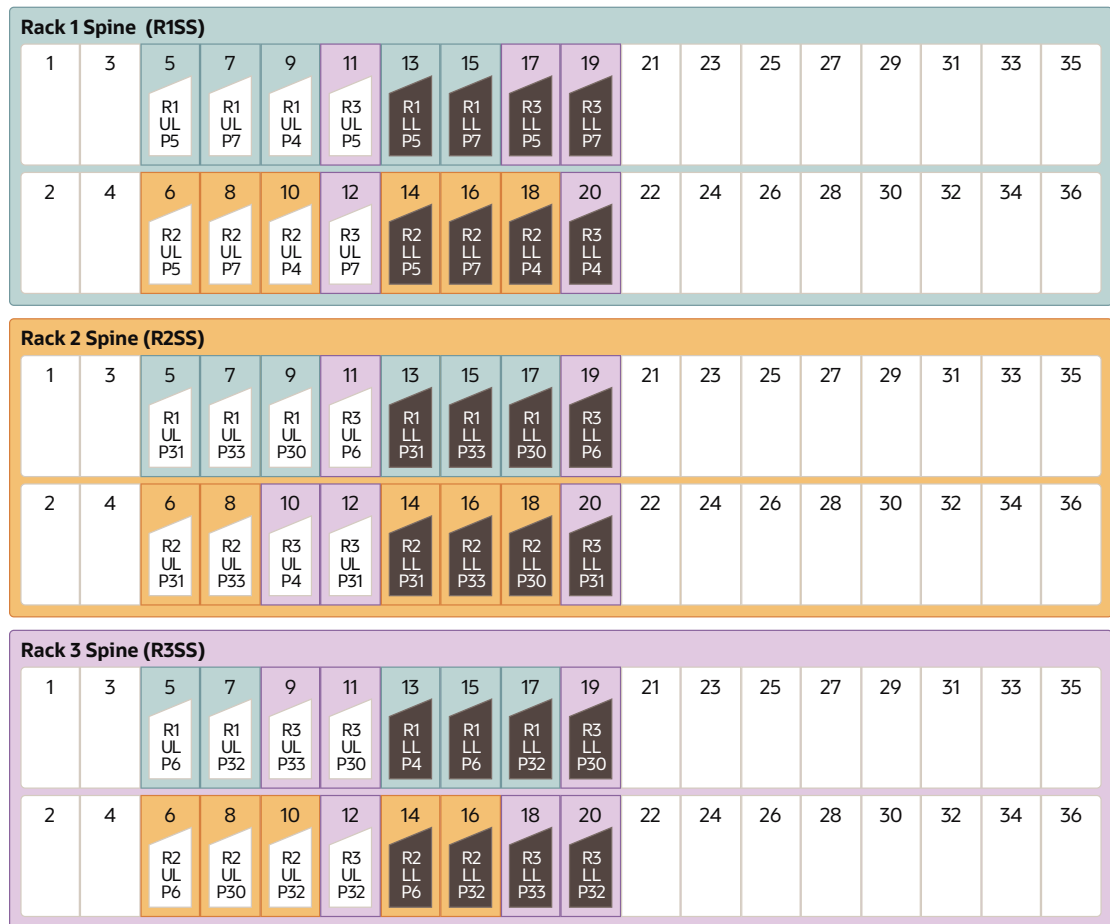
This section provides the cabling tables to connect three X8M racks together using RoCE Network Fabric.

 **Note:**

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the three spine switches (Rack1-spine, Rack2-spine, and Rack3-spine) when cabling three racks together:



The following table describes the cable connections for the first spine switch (R1-SS) when cabling three racks together:

Table 5-3 Leaf Switch Connections for the First Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5 R1-UL-P7 to R1-SS-P7 R1-UL-P4 to R1-SS-P9	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5 R1-UL-P33 to R2-SS-P7 R1-UL-P30 to R2-SS-P9	5 meters
R1 UL to Rack 3	R1-UL-P6 to R3-SS-P5 R1-UL-P32 to R3-SS-P7	5 meters
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13 R1-LL-P7 to R1-SS-P15	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13 R1-LL-P33 to R2-SS-P15 R1-LL-P30 to R2-SS-P17	5 meters

Table 5-3 (Cont.) Leaf Switch Connections for the First Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R1 LL to Rack 3	R1-LL-P4 to R3-SS-P13 R1-LL-P6 to R3-SS-P15 R1-LL-P32 to R3-SS-P17	5 meters

The following table describes the cable connections for the second spine switch (R2-SS) when cabling three racks together:

Table 5-4 Leaf Switch Connections for the Second Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6 R2-UL-P7 to R1-SS-P8 R2-UL-P4 to R1-SS-P10	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6 R2-UL-P33 to R2-SS-P8	3 meters
R2 UL to Rack 3	R2-UL-P6 to R3-SS-P6 R2-UL-P30 to R3-SS-P8 R2-UL-P32 to R3-SS-P10	5 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14 R2-LL-P7 to R1-SS-P16 R2-LL-P4 to R1-SS-P18	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14 R2-LL-P33 to R2-SS-P16 R2-LL-P30 to R2-SS-P18	3 meters
R2 LL to Rack 3	R2-LL-P6 to R3-SS-P14 R2-LL-P32 to R3-SS-P16	5 meters

The following table describes the cable connections for the third spine switch (R3-SS) when cabling three racks together:

Table 5-5 Leaf Switch Connections for the Third Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 1	R3-UL-P5 to R1-SS-P11 R3-UL-P7 to R1-SS-P12	5 meters
R3 UL to Rack 2	R3-UL-P4 to R2-SS-P10 R3-UL-P6 to R2-SS-P11 R3-UL-P31 to R2-SS-P12	5 meters
R3 UL within Rack 3	R3-UL-P33 to R3-SS-P9 R3-UL-P30 to R3-SS-P11 R3-UL-P32 to R3-SS-P12	3 meters

Table 5-5 (Cont.) Leaf Switch Connections for the Third Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R3 LL to Rack 1	R3-LL-P5 to R1-SS-P17 R3-LL-P7 to R1-SS-P19 R3-LL-P4 to R1-SS-P20	5 meters
R3 LL to Rack 2	R3-LL-P6 to R2-SS-P19 R3-LL-P31 to R2-SS-P20	5 meters
R3 LL within Rack 3	R3-LL-P33 to R3-SS-P18 R3-LL-P30 to R3-SS-P19 R3-LL-P32 to R3-SS-P20	3 meters

5.4 Four-Rack Cabling for X8M Racks

This section provides the cabling tables to connect four X8M racks together, all of which use RoCE Network Fabric.

 **Note:**

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the four spine switches (Rack1-spine, Rack2-spine, Rack3-spine, and Rack4-spine) when cabling two racks together:

Rack 1 Spine (R1SS)																	
1	3	5 R1 UL P5	7 R1 UL P7	9 R4 UL P5	11 R3 UL P5	13 R1 LL P5	15 R1 LL P7	17 R4 LL P5	19 R3 LL P7	21	23	25	27	29	31	33	35
2	4	6 R2 UL P5	8 R2 UL P7	10 R4 UL P7	12 R3 UL P7	14 R2 LL P5	16 R2 LL P7	18 R4 LL P7	20 R3 LL P4	22	24	26	28	30	32	34	36

Rack 2 Spine (R2SS)																	
1	3	5 R1 UL P31	7 R1 UL P33	9 R4 UL P4	11 R3 UL P6	13 R1 LL P31	15 R1 LL P33	17 R4 LL P4	19 R3 LL P6	21	23	25	27	29	31	33	35
2	4	6 R2 UL P31	8 R2 UL P33	10 R4 UL P6	12 R3 UL P31	14 R2 LL P31	16 R2 LL P33	18 R4 LL P6	20 R3 LL P31	22	24	26	28	30	32	34	36

Rack 3 Spine (R3SS)																	
1	3	5 R1 UL P6	7 R1 UL P32	9 R4 UL P31	11 R3 UL P30	13 R1 LL P4	15 R1 LL P6	17 R4 LL P31	19 R3 LL P30	21	23	25	27	29	31	33	35
2	4	6 R2 UL P6	8 R2 UL P30	10 R4 UL P33	12 R3 UL P32	14 R2 LL P6	16 R2 LL P32	18 R4 LL P33	20 R3 LL P32	22	24	26	28	30	32	34	36

Rack 4 Spine (R4SS)																	
1	3	5 R1 UL P4	7 R1 UL P30	9 R4 UL P30	11 R3 UL P4	13 R1 LL P30	15 R1 LL P32	17 R4 LL P30	19 R3 LL P5	21	23	25	27	29	31	33	35
2	4	6 R2 UL P4	8 R2 UL P32	10 R4 UL P32	12 R3 UL P32	14 R2 LL P4	16 R2 LL P30	18 R4 LL P32	20 R3 LL P33	22	24	26	28	30	32	34	36

The following table describes the cable connections for the first spine switch (R1-SS) when cabling four racks together:

Table 5-6 Leaf Switch Connections for the First Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5 R1-UL-P7 to R1-SS-P7	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5 R1-UL-P33 to R2-SS-P7	5 meters
R1 UL to Rack 3	R1-UL-P6 to R3-SS-P5 R1-UL-P32 to R3-SS-P7	5 meters

Table 5-6 (Cont.) Leaf Switch Connections for the First Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R1 UL to Rack 4	R1-UL-P4 to R4-SS-P5 R1-UL-P30 to R4-SS-P7	10 meters
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13 R1-LL-P7 to R1-SS-P15	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13 R1-LL-P33 to R2-SS-P15	5 meters
R1 LL to Rack 3	R1-LL-P4 to R3-SS-P13 R1-LL-P6 to R3-SS-P15	5 meters
R1 LL to Rack 4	R1-LL-P30 to R4-SS-P13 R1-LL-P32 to R4-SS-P15	10 meters

The following table describes the cable connections for the second spine switch (R2-SS) when cabling four full racks together:

Table 5-7 Leaf Switch Connections for the Second Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6 R2-UL-P7 to R1-SS-P8	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6 R2-UL-P33 to R2-SS-P8	3 meters
R2 UL to Rack 3	R2-UL-P6 to R3-SS-P6 R2-UL-P30 to R3-SS-P8	5 meters
R2 UL to Rack 4	R2-UL-P4 to R4-SS-P6 R2-UL-P32 to R4-SS-P8	5 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14 R2-LL-P7 to R1-SS-P16	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14 R2-LL-P33 to R2-SS-P16	3 meters
R2 LL to Rack 3	R2-LL-P6 to R3-SS-P14 R2-LL-P32 to R3-SS-P16	5 meters
R2 LL to Rack 4	R2-LL-P4 to R4-SS-P14 R2-LL-P30 to R4-SS-P16	5 meters

The following table describes the cable connections for the third spine switch (R3-SS) when cabling four full racks together:

Table 5-8 Leaf Switch Connections for the Third Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 1	R3-UL-P5 to R1-SS-P11 R3-UL-P7 to R1-SS-P12	5 meters
R3 UL to Rack 2	R3-UL-P6 to R2-SS-P11 R3-UL-P31 to R2-SS-P12	5 meters
R3 UL within Rack 3	R3-UL-P30 to R3-SS-P11 R3-UL-P32 to R3-SS-P12	3 meters
R3 UL to Rack 4	R3-UL-P4 to R4-SS-P11 R3-UL-P33 to R4-SS-P12	5 meters
R3 LL to Rack 1	R3-LL-P7 to R1-SS-P19 R3-LL-P4 to R1-SS-P20	5 meters
R3 LL to Rack 2	R3-LL-P6 to R2-SS-P19 R3-LL-P31 to R2-SS-P20	5 meters
R3 LL within Rack 3	R3-LL-P30 to R3-SS-P19 R3-LL-P32 to R3-SS-P20	3 meters
R3 LL to Rack 4	R3-LL-P5 to R4-SS-P19 R3-LL-P33 to R4-SS-P20	5 meters

The following table describes the cable connections for the fourth spine switch (R4-SS) when cabling four full racks together:

Table 5-9 Leaf Switch Connections for the Fourth Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R4 UL to Rack 1	R4-UL-P5 to R1-SS-P9 R4-UL-P7 to R1-SS-P10	10 meters
R4 UL to Rack 2	R4-UL-P4 to R2-SS-P9 R4-UL-P6 to R2-SS-P10	5 meters
R4 UL to Rack 3	R4-UL-P31 to R3-SS-P9 R4-UL-P33 to R3-SS-P10	5 meters
R4 UL within Rack 4	R4-UL-P30 to R4-SS-P9 R4-UL-P32 to R4-SS-P10	3 meters
R4 LL to Rack 1	R4-LL-P5 to R1-SS-P17 R4-LL-P7 to R1-SS-P18	10 meters
R4 LL to Rack 2	R4-LL-P4 to R2-SS-P17 R4-LL-P6 to R2-SS-P18	5 meters
R4 LL to Rack 3	R4-LL-P31 to R3-SS-P17 R4-LL-P33 to R3-SS-P18	5 meters
R4 LL within Rack 4	R4-LL-P30 to R4-SS-P17 R4-LL-P32 to R4-SS-P18	3 meters

5.5 Five-Rack Cabling for X8M Racks

This section provides the cabling tables to connect five (5) X8M racks together, all of which use RoCE Network Fabric.

Note:

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the five spine switches when cabling five racks together:

Rack 1 Spine (R1SS)																	
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
		R1 UL P5	R1 UL P7	R5 UL P7	R3 UL P5	R1 LL P5	R5 LL P5	R4 LL P5	R3 LL P7								
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
		R2 UL P5	R5 UL P5	R4 UL P7	R3 UL P7	R2 LL P5	R2 LL P7	R5 LL P7	R3 LL P4								

Rack 2 Spine (R2SS)																	
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
		R1 UL P31	R1 UL P33	R4 UL P4	R3 UL P6	R1 LL P31	R1 LL P33	R4 LL P4	R3 LL P6								
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
		R2 UL P31	R5 UL P4	R4 UL P6	R3 UL P31	R2 LL P31	R5 LL P4	R5 LL P6	R3 LL P31								

Rack 3 Spine (R3SS)																	
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
		R1 UL P6	R1 UL P32	R4 UL P31	R3 UL P30	R1 LL P4	R1 LL P6	R4 LL P31	R3 LL P30								
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
		R2 UL P6	R2 UL P30	R4 UL P33	R5 UL P6	R2 LL P6	R5 LL P31	R4 LL P33	R3 LL P32								

Rack 4 Spine (R4SS)																	
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
		R1 UL P4	R5 UL P31	R4 UL P30	R3 UL P4	R1 LL P30	R1 LL P32	R4 LL P30	R5 LL P33								
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
		R2 UL P4	R2 UL P32	R4 UL P32	R5 UL P33	R2 LL P4	R2 LL P30	R4 LL P32	R3 LL P33								

Rack 5 Spine (R5SS)																	
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
		R1 UL P30	R5 UL P30	R5 UL P32	R3 UL P32	R1 LL P7	R5 LL P30	R4 LL P7	R5 LL P32								
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
		R2 UL P7	R2 UL P33	R4 UL P5	R3 UL P33	R2 LL P33	R2 LL P32	R4 LL P6	R3 LL P5								

The following table describes the cable connections for the first spine switch (R1-SS) when cabling five racks together:

Table 5-10 Leaf Switch Connections for the First Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5 R1-UL-P7 to R1-SS-P7	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5 R1-UL-P33 to R2-SS-P7	5 meters
R1 UL to Rack 3	R1-UL-P6 to R3-SS-P5 R1-UL-P32 to R3-SS-P7	5 meters
R1 UL to Rack 4	R1-UL-P4 to R4-SS-P5	10 meters
R1 UL to Rack 5	R1-UL-P30 to R5-SS-P5	10 meters
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13 R1-LL-P33 to R2-SS-P15	5 meters
R1 LL to Rack 3	R1-LL-P4 to R3-SS-P13 R1-LL-P6 to R3-SS-P15	5 meters
R1 LL to Rack 4	R1-LL-P30 to R4-SS-P13 R1-LL-P32 to R4-SS-P15	10 meters
R1 LL to Rack 5	R1-LL-P7 to R5-SS-P13	10 meters

The following table describes the cable connections for the second spine switch (R2-SS) when cabling five full racks together:

Table 5-11 Leaf Switch Connections for the Second Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6	3 meters
R2 UL to Rack 3	R2-UL-P6 to R3-SS-P6 R2-UL-P30 to R3-SS-P8	5 meters
R2 UL to Rack 4	R2-UL-P4 to R4-SS-P6 R2-UL-P32 to R4-SS-P8	5 meters
R2 UL to Rack 5	R2-UL-P7 to R5-SS-P6 R2-UL-P33 to R5-SS-P8	10 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14 R2-LL-P7 to R1-SS-P16	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14	3 meters
R2 LL to Rack 3	R2-LL-P6 to R3-SS-P14	5 meters
R2 LL to Rack 4	R2-LL-P4 to R4-SS-P14 R2-LL-P30 to R4-SS-P16	5 meters
R2 LL to Rack 5	R2-LL-P33 to R5-SS-P14 R2-LL-P32 to R5-SS-P16	10 meters

The following table describes the cable connections for the third spine switch (R3-SS) when cabling five full racks together:

Table 5-12 Leaf Switch Connections for the Third Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 1	R3-UL-P5 to R1-SS-P11 R3-UL-P7 to R1-SS-P12	5 meters
R3 UL to Rack 2	R3-UL-P6 to R2-SS-P11 R3-UL-P31 to R2-SS-P12	5 meters
R3 UL within Rack 3	R3-UL-P30 to R3-SS-P11	3 meters
R3 UL to Rack 4	R3-UL-P4 to R4-SS-P11	5 meters
R3 UL to Rack 5	R3-UL-P32 to R5-SS-P11 R3-UL-P33 to R5-SS-P12	5 meters
R3 LL to Rack 1	R3-LL-P7 to R1-SS-P19 R3-LL-P4 to R1-SS-P20	5 meters
R3 LL to Rack 2	R3-LL-P6 to R2-SS-P19 R3-LL-P31 to R2-SS-P20	5 meters
R3 LL within Rack 3	R3-LL-P30 to R3-SS-P19 R3-LL-P32 to R3-SS-P20	3 meters
R3 LL to Rack 4	R3-LL-P33 to R4-SS-P20	5 meters
R3 LL to Rack 5	R3-LL-P5 to R5-SS-P20	5 meters

The following table describes the cable connections for the fourth spine switch (R4-SS) when cabling five full racks together:

Table 5-13 Leaf Switch Connections for the Fourth Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R4 UL to Rack 1	R4-UL-P7 to R1-SS-P10	10 meters
R4 UL to Rack 2	R4-UL-P4 to R2-SS-P9 R4-UL-P6 to R2-SS-P10	5 meters
R4 UL to Rack 3	R4-UL-P31 to R3-SS-P9 R4-UL-P33 to R3-SS-P10	5 meters
R4 UL within Rack 4	R4-UL-P30 to R4-SS-P9 R4-UL-P32 to R4-SS-P10	3 meters
R4 UL to Rack 5	R4-UL-P5 to R5-SS-P10	5 meters
R4 LL to Rack 1	R4-LL-P5 to R1-SS-P17	10 meters
R4 LL to Rack 2	R4-LL-P4 to R2-SS-P17	5 meters
R4 LL to Rack 3	R4-LL-P31 to R3-SS-P17 R4-LL-P33 to R3-SS-P18	5 meters
R4 LL within Rack 4	R4-LL-P30 to R4-SS-P17 R4-LL-P32 to R4-SS-P18	3 meters

Table 5-13 (Cont.) Leaf Switch Connections for the Fourth Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R4 LL to Rack 5	R4-LL-P7 to R5-SS-P17 R4-LL-P6 to R5-SS-P18	5 meters

The following table describes the cable connections for the fifth spine switch (R5-SS) when cabling five full racks together:

Table 5-14 Leaf Switch Connections for the Fifth Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R5 UL to Rack 1	R5-UL-P5 to R1-SS-P8 R5-UL-P7 to R1-SS-P9	10 meters
R5 UL to Rack 2	R5-UL-P4 to R2-SS-P8	10 meters
R5 UL to Rack 3	R5-UL-P6 to R3-SS-P12	5 meters
R5 UL to Rack 4	R5-UL-P31 to R4-SS-P7 R5-UL-P33 to R4-SS-P12	5 meters
R5 UL within Rack 5	R5-UL-P30 to R5-SS-P7 R5-UL-P32 to R5-SS-P9	3 meters
R5 LL to Rack 1	R5-LL-P5 to R1-SS-P15 R5-LL-P7 to R1-SS-P18	10 meters
R5 LL to Rack 2	R5-LL-P4 to R2-SS-P16 R5-LL-P6 to R2-SS-P18	10 meters
R5 LL to Rack 3	R5-LL-P31 to R3-SS-P16	5 meters
R5 LL to Rack 4	R5-LL-P33 to R4-SS-P19	5 meters
R5 LL within Rack 5	R5-LL-P30 to R5-SS-P15 R5-LL-P32 to R5-SS-P19	3 meters

5.6 Six-Rack Cabling for X8M Racks

This section provides the cabling tables to connect six (6) X8M racks together, all of which use RoCE Network Fabric.

 **Note:**

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following illustration shows the cable connections for the 6 spine switches when cabling six racks together:

Rack 1 Spine (R1SS)

1	3	5 R1 UL P5	7 R1 UL P7	9 R6 UL P5	11 R3 UL P5	13 R1 LL P5	15 R6 LL P7	17 R4 LL P5	19 R3 LL P7	21	23	25	27	29	31	33	35
2	4	6 R2 UL P5	8 R5 UL P5	10 R4 UL P7	12 R6 LL P5	14 R2 LL P5	16 R6 UL P7	18 R5 LL P7	20 R3 LL P4	22	24	26	28	30	32	34	36

Rack 2 Spine (R2SS)

1	3	5 R1 UL P31	7 R1 UL P33	9 R4 UL P4	11 R3 UL P6	13 R1 LL P31	15 R1 LL P33	17 R4 LL P4	19 R6 LL P6	21	23	25	27	29	31	33	35
2	4	6 R2 UL P31	8 R5 UL P4	10 R4 UL P6	12 R6 LL P4	14 R2 LL P31	16 R6 UL P4	18 R5 LL P6	20 R3 LL P31	22	24	26	28	30	32	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P6	7 R6 LL P31	9 R4 UL P31	11 R3 UL P30	13 R1 LL P4	15 R1 LL P6	17 R4 LL P31	19 R6 UL P6	21	23	25	27	29	31	33	35
2	4	6 R2 UL P6	8 R2 UL P30	10 R4 UL P33	12 R5 UL P6	14 R2 LL P6	16 R5 LL P31	18 R4 LL P33	20 R3 LL P32	22	24	26	28	30	32	34	36

Rack 4 Spine (R4SS)

1	3	5 R1 UL P4	7 R5 UL P31	9 R6 LL P33	11 R3 UL P4	13 R1 LL P30	15 R6 UL P31	17 R4 LL P30	19 R5 LL P33	21	23	25	27	29	31	33	35
2	4	6 R2 UL P4	8 R2 UL P32	10 R4 UL P32	12 R5 UL P33	14 R2 LL P4	16 R2 LL P30	18 R4 LL P32	20 R3 LL P33	22	24	26	28	30	32	34	36

Rack 5 Spine (R5SS)

1	3	5 R1 UL P30	7 R5 UL P30	9 R5 UL P32	11 R3 UL P32	13 R1 LL P7	15 R5 LL P30	17 R4 LL P7	19 R5 LL P32	21	23	25	27	29	31	33	35
2	4	6 R2 UL P7	8 R6 LL P30	10 R4 UL P5	12 R3 UL P33	14 R2 LL P33	16 R2 LL P32	18 R6 UL P33	20 R3 LL P5	22	24	26	28	30	32	34	36

Rack 6 Spine (R6SS)

1	3	5 R1 UL P32	7 R5 UL P7	9 R5 LL P5	11 R3 UL P7	13 R1 LL P32	15 R5 LL P4	17 R4 LL P6	19 R3 LL P6	21	23	25	27	29	31	33	35
2	4	6 R2 UL P33	8 R6 LL P32	10 R4 UL P30	12 R3 UL P31	14 R2 LL P7	16 R6 UL P30	18 R6 UL P32	20 R3 LL P30	22	24	26	28	30	32	34	36

The following table describes the cable connections for the first spine switch (R1-SS) when cabling six racks together:

Table 5-15 Leaf Switch Connections for the First Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5 R1-UL-P7 to R1-SS-P7	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5 R1-UL-P33 to R2-SS-P7	5 meters
R1 UL to Rack 3	R1-UL-P6 to R3-SS-P5	5 meters
R1 UL to Rack 4	R1-UL-P4 to R4-SS-P5	10 meters
R1 UL to Rack 5	R1-UL-P30 to R5-SS-P5	10 meters
R1 UL to Rack 6	R1-UL-P32 to R6-SS-P5	10 meters
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13 R1-LL-P33 to R2-SS-P15	5 meters
R1 LL to Rack 3	R1-LL-P4 to R3-SS-P13 R1-LL-P6 to R3-SS-P15	5 meters
R1 LL to Rack 4	R1-LL-P30 to R4-SS-P13	10 meters
R1 LL to Rack 5	R1-LL-P7 to R5-SS-P13	10 meters
R1 LL to Rack 6	R1-LL-P32 to R6-SS-P13	10 meters

The following table describes the cable connections for the second spine switch (R2-SS) when cabling six full racks together:

Table 5-16 Leaf Switch Connections for the Second Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6	3 meters
R2 UL to Rack 3	R2-UL-P6 to R3-SS-P6 R2-UL-P30 to R3-SS-P8	5 meters
R2 UL to Rack 4	R2-UL-P4 to R4-SS-P6 R2-UL-P32 to R4-SS-P8	5 meters
R2 UL to Rack 5	R2-UL-P7 to R5-SS-P6	10 meters
R2 UL to Rack 6	R2-UL-P33 to R6-SS-P6	10 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14	3 meters
R2 LL to Rack 3	R2-LL-P6 to R3-SS-P14	5 meters
R2 LL to Rack 4	R2-LL-P4 to R4-SS-P14 R2-LL-P30 to R4-SS-P16	5 meters
R2 LL to Rack 5	R2-LL-P33 to R5-SS-P14 R2-LL-P32 to R5-SS-P16	10 meters

Table 5-16 (Cont.) Leaf Switch Connections for the Second Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R2 LL to Rack 6	R2-LL-P7 to R6-SS-P14	10 meters

The following table describes the cable connections for the third spine switch (R3-SS) when cabling six full racks together:

Table 5-17 Leaf Switch Connections for the Third Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 1	R3-UL-P5 to R1-SS-P11	5 meters
R3 UL to Rack 2	R3-UL-P6 to R2-SS-P11	5 meters
R3 UL within Rack 3	R3-UL-P30 to R3-SS-P11	3 meters
R3 UL to Rack 4	R3-UL-P4 to R4-SS-P11	5 meters
R3 UL to Rack 5	R3-UL-P32 to R5-SS-P11 R3-UL-P33 to R5-SS-P12	5 meters
R3 UL to Rack 6	R3-UL-P7 to R6-SS-P11 R3-UL-P31 to R6-SS-P12	10 meters
R3 LL to Rack 1	R3-LL-P7 to R1-SS-P19 R3-LL-P4 to R1-SS-P20	5 meters
R3 LL to Rack 2	R3-LL-P31 to R2-SS-P20	5 meters
R3 LL within Rack 3	R3-LL-P32 to R3-SS-P20	3 meters
R3 LL to Rack 4	R3-LL-P33 to R4-SS-P20	5 meters
R3 LL to Rack 5	R3-LL-P5 to R5-SS-P20	5 meters
R3 LL to Rack 6	R3-LL-P6 to R6-SS-P19 R3-LL-P30 to R6-SS-P20	10 meters

The following table describes the cable connections for the fourth spine switch (R4-SS) when cabling six full racks together:

Table 5-18 Leaf Switch Connections for the Fourth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R4 UL to Rack 1	R4-UL-P7 to R1-SS-P10	10 meters
R4 UL to Rack 2	R4-UL-P4 to R2-SS-P9 R4-UL-P6 to R2-SS-P10	5 meters
R4 UL to Rack 3	R4-UL-P31 to R3-SS-P9 R4-UL-P33 to R3-SS-P10	5 meters
R4 UL within Rack 4	R4-UL-P32 to R4-SS-P10	3 meters
R4 UL to Rack 5	R4-UL-P5 to R5-SS-P10	5 meters
R4 UL to Rack 6	R4-UL-P30 to R6-SS-P10	5 meters

Table 5-18 (Cont.) Leaf Switch Connections for the Fourth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R4 LL to Rack 1	R4-LL-P5 to R1-SS-P17	10 meters
R4 LL to Rack 2	R4-LL-P4 to R2-SS-P17	5 meters
R4 LL to Rack 3	R4-LL-P31 to R3-SS-P17 R4-LL-P33 to R3-SS-P18	5 meters
R4 LL within Rack 4	R4-LL-P30 to R4-SS-P17 R4-LL-P32 to R4-SS-P18	3 meters
R4 LL to Rack 5	R4-LL-P7 to R5-SS-P17	5 meters
R4 LL to Rack 6	R4-LL-P6 to R6-SS-P17	5 meters

The following table describes the cable connections for the fifth spine switch (R5-SS) when cabling six full racks together:

Table 5-19 Leaf Switch Connections for the Fifth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R5 UL to Rack 1	R5-UL-P5 to R1-SS-P8	10 meters
R5 UL to Rack 2	R5-UL-P4 to R2-SS-P8	10 meters
R5 UL to Rack 3	R5-UL-P6 to R3-SS-P12	5 meters
R5 UL to Rack 4	R5-UL-P31 to R4-SS-P7 R5-UL-P33 to R4-SS-P12	5 meters
R5 UL within Rack 5	R5-UL-P30 to R5-SS-P7 R5-UL-P32 to R5-SS-P9	3 meters
R5 UL to Rack 6	R5-UL-P7 to R6-SS-P7	5 meters
R5 LL to Rack 1	R5-LL-P7 to R1-SS-P18	10 meters
R5 LL to Rack 2	R5-LL-P6 to R2-SS-P18	10 meters
R5 LL to Rack 3	R5-LL-P31 to R3-SS-P16	5 meters
R5 LL to Rack 4	R5-LL-P33 to R4-SS-P19	5 meters
R5 LL within Rack 5	R5-LL-P30 to R5-SS-P15 R5-LL-P32 to R5-SS-P19	3 meters
R5 LL to Rack 6	R5-LL-P5 to R6-SS-P9 R5-LL-P4 to R6-SS-P15	5 meters

The following table describes the cable connections for the sixth spine switch (R6-SS) when cabling six full racks together:

Table 5-20 Leaf Switch Connections for the Sixth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R6 UL to Rack 1	R6-UL-P5 to R1-SS-P9 R6-UL-P7 to R1-SS-P16	10 meters

Table 5-20 (Cont.) Leaf Switch Connections for the Sixth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R6 UL to Rack 2	R6-UL-P4 to R2-SS-P16	10 meters
R6 UL to Rack 3	R6-UL-P6 to R3-SS-P19	10 meters
R6 UL to Rack 4	R6-UL-P31 to R4-SS-P15	5 meters
R6 UL to Rack 5	R6-UL-P33 to R5-SS-P18	5 meters
R6 UL within Rack 6	R6-UL-P30 to R6-SS-P16 R6-UL-P32 to R6-SS-P18	3 meters
R6 LL to Rack 1	R6-LL-P5 to R1-SS-P12 R6-LL-P7 to R1-SS-P15	10 meters
R6 LL to Rack 2	R6-LL-P4 to R2-SS-P12 R6-LL-P6 to R2-SS-P19	10 meters
R6 LL to Rack 3	R6-LL-P31 to R3-SS-P7	10 meters
R6 LL to Rack 4	R6-LL-P33 to R4-SS-P9	5 meters
R6 LL to Rack 5	R6-LL-P30 to R5-SS-P8	5 meters
R6 LL within Rack 6	R6-LL-P32 to R6-SS-P8	3 meters

5.7 Seven-Rack Cabling for X8M Racks

This section provides the cabling tables to connect seven (7) X8M racks together, all of which use RoCE Network Fabric.

 **Note:**

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following diagrams show the cable connections for the 7 spine switches when cabling seven racks together:

Rack 1 Spine (R1SS)

1	3	5 R1 UL P5	7 R1 UL P7	9 R7 UL P5	11 R3 UL P5	13 R1 LL P5	15 R6 LL P7	17 R4 LL P5	19 R7 LL P5	21	23	25	27	29	31	33	35
2	4	6 R2 UL P5	8 R5 UL P5	10 R4 UL P7	12 R6 LL P5	14 R2 LL P5	16 R6 UL P7	18 R5 LL P7	20 R3 LL P4	22	24	26	28	30	32	34	36

Rack 2 Spine (R2SS)

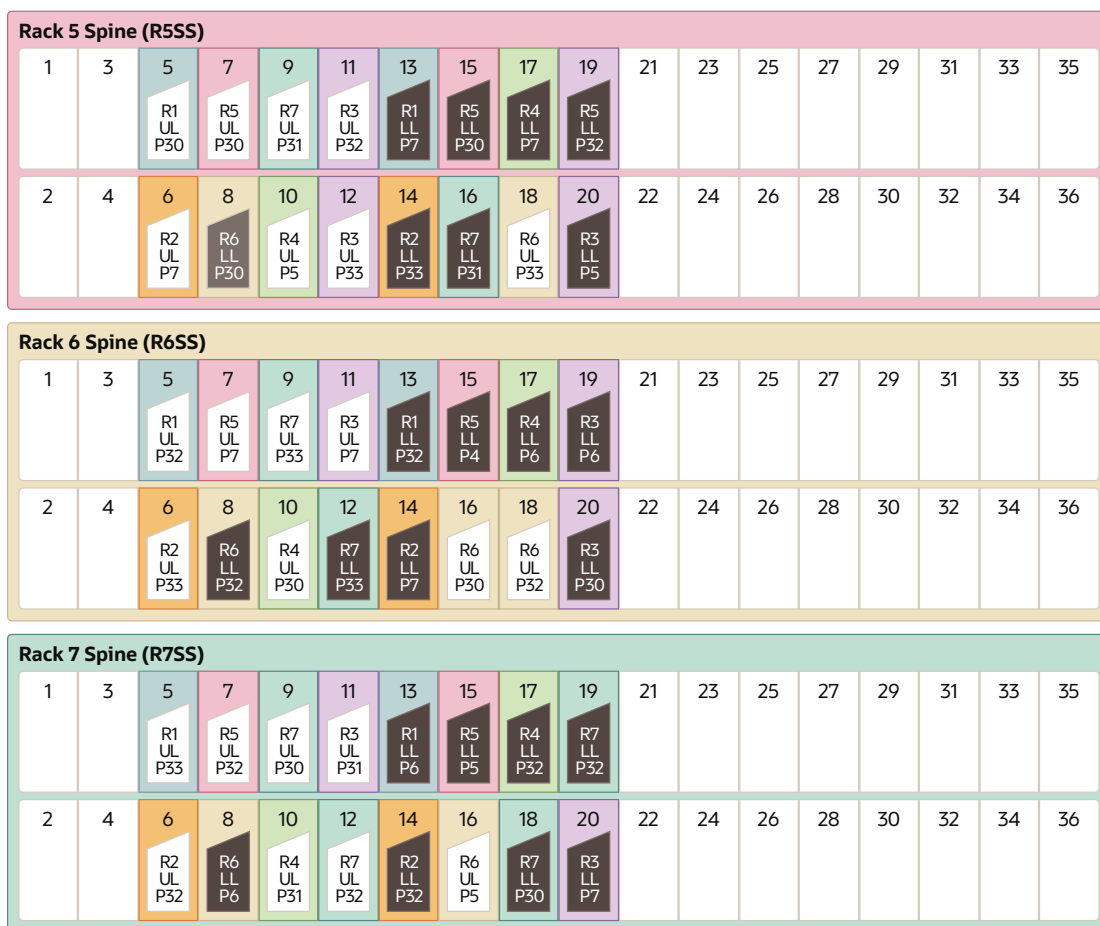
1	3	5 R1 UL P31	7 R7 UL P7	9 R4 UL P4	11 R3 UL P6	13 R1 LL P31	15 R1 LL P33	17 R4 LL P4	19 R7 LL P7	21	23	25	27	29	31	33	35
2	4	6 R2 UL P31	8 R5 UL P4	10 R4 UL P6	12 R6 LL P4	14 R2 LL P31	16 R6 UL P4	18 R5 LL P6	20 R3 LL P31	22	24	26	28	30	32	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P6	7 R6 LL P31	9 R7 UL P4	11 R3 UL P30	13 R1 LL P4	15 R7 LL P4	17 R4 LL P31	19 R6 UL P6	21	23	25	27	29	31	33	35
2	4	6 R2 UL P6	8 R2 UL P30	10 R4 UL P33	12 R5 UL P6	14 R2 LL P6	16 R5 LL P31	18 R4 LL P33	20 R3 LL P32	22	24	26	28	30	32	34	36

Rack 4 Spine (R4SS)

1	3	5 R1 UL P4	7 R5 UL P31	9 R6 LL P33	11 R3 UL P4	13 R1 LL P30	15 R6 UL P31	17 R4 LL P30	19 R5 LL P33	21	23	25	27	29	31	33	35
2	4	6 R2 UL P4	8 R7 UL P6	10 R4 UL P32	12 R5 UL P33	14 R2 LL P4	16 R2 LL P30	18 R7 LL P6	20 R3 LL P33	22	24	26	28	30	32	34	36



The following table describes the cable connections for the first spine switch (R1-SS) when cabling seven racks together:

Table 5-21 Leaf Switch Connections for the First Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5 R1-UL-P7 to R1-SS-P7	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5	5 meters
R1 UL to Rack 3	R1-UL-P6 to R3-SS-P5	5 meters
R1 UL to Rack 4	R1-UL-P4 to R4-SS-P5	10 meters
R1 UL to Rack 5	R1-UL-P30 to R5-SS-P5	10 meters
R1 UL to Rack 6	R1-UL-P32 to R6-SS-P5	10 meters
R1 UL to Rack 7	R1-UL-P33 to R7-SS-P5	10 meters
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13 R1-LL-P33 to R2-SS-P15	5 meters
R1 LL to Rack 3	R1-LL-P4 to R3-SS-P13	5 meters
R1 LL to Rack 4	R1-LL-P30 to R4-SS-P13	10 meters

Table 5-21 (Cont.) Leaf Switch Connections for the First Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R1 LL to Rack 5	R1-LL-P7 to R5-SS-P13	10 meters
R1 LL to Rack 6	R1-LL-P32 to R6-SS-P13	10 meters
R1 LL to Rack 7	R1-LL-P6 to R7-SS-P13	10 meters

The following table describes the cable connections for the second spine switch (R2-SS) when cabling seven full racks together:

Table 5-22 Leaf Switch Connections for the Second Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6	3 meters
R2 UL to Rack 3	R2-UL-P6 to R3-SS-P6 R2-UL-P30 to R3-SS-P8	5 meters
R2 UL to Rack 4	R2-UL-P4 to R4-SS-P6	5 meters
R2 UL to Rack 5	R2-UL-P7 to R5-SS-P6	10 meters
R2 UL to Rack 6	R2-UL-P33 to R6-SS-P6	10 meters
R2 UL to Rack 7	R2-UL-P32 to R7-SS-P6	10 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14	3 meters
R2 LL to Rack 3	R2-LL-P6 to R3-SS-P14	5 meters
R2 LL to Rack 4	R2-LL-P4 to R4-SS-P14 R2-LL-P30 to R4-SS-P16	5 meters
R2 LL to Rack 5	R2-LL-P33 to R5-SS-P14	10 meters
R2 LL to Rack 6	R2-LL-P7 to R6-SS-P14	10 meters
R2 LL to Rack 7	R2-LL-P32 to R7-SS-P14	10 meters

The following table describes the cable connections for the third spine switch (R3-SS) when cabling seven full racks together:

Table 5-23 Leaf Switch Connections for the Third Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 1	R3-UL-P5 to R1-SS-P11	5 meters
R3 UL to Rack 2	R3-UL-P6 to R2-SS-P11	5 meters
R3 UL within Rack 3	R3-UL-P30 to R3-SS-P11	3 meters
R3 UL to Rack 4	R3-UL-P4 to R4-SS-P11	5 meters

Table 5-23 (Cont.) Leaf Switch Connections for the Third Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 5	R3-UL-P32 to R5-SS-P11 R3-UL-P33 to R5-SS-P12	5 meters
R3 UL to Rack 6	R3-UL-P7 to R6-SS-P11	10 meters
R3 UL to Rack 7	R3-UL-P31 to R7-SS-P11	10 meters
R3 LL to Rack 1	R3-LL-P4 to R1-SS-P20	5 meters
R3 LL to Rack 2	R3-LL-P31 to R2-SS-P20	5 meters
R3 LL within Rack 3	R3-LL-P32 to R3-SS-P20	3 meters
R3 LL to Rack 4	R3-LL-P33 to R4-SS-P20	5 meters
R3 LL to Rack 5	R3-LL-P5 to R5-SS-P20	5 meters
R3 LL to Rack 6	R3-LL-P6 to R6-SS-P19 R3-LL-P30 to R6-SS-P20	10 meters
R3 LL to Rack 7	R3-LL-P7 to R7-SS-P20	10 meters

The following table describes the cable connections for the fourth spine switch (R4-SS) when cabling seven full racks together:

Table 5-24 Leaf Switch Connections for the Fourth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R4 UL to Rack 1	R4-UL-P7 to R1-SS-P10	10 meters
R4 UL to Rack 2	R4-UL-P4 to R2-SS-P9 R4-UL-P6 to R2-SS-P10	5 meters
R4 UL to Rack 3	R4-UL-P33 to R3-SS-P10	5 meters
R4 UL within Rack 4	R4-UL-P32 to R4-SS-P10	3 meters
R4 UL to Rack 5	R4-UL-P5 to R5-SS-P10	5 meters
R4 UL to Rack 6	R4-UL-P30 to R6-SS-P10	5 meters
R4 UL to Rack 7	R4-UL-P31 to R7-SS-P10	10 meters
R4 LL to Rack 1	R4-LL-P5 to R1-SS-P17	10 meters
R4 LL to Rack 2	R4-LL-P4 to R2-SS-P17	5 meters
R4 LL to Rack 3	R4-LL-P31 to R3-SS-P17 R4-LL-P33 to R3-SS-P18	5 meters
R4 LL within Rack 4	R4-LL-P30 to R4-SS-P17	3 meters
R4 LL to Rack 5	R4-LL-P7 to R5-SS-P17	5 meters
R4 LL to Rack 6	R4-LL-P6 to R6-SS-P17	5 meters
R4 LL to Rack 7	R4-LL-P32 to R7-SS-P17	10 meters

The following table describes the cable connections for the fifth spine switch (R5-SS) when cabling seven full racks together:

Table 5-25 Leaf Switch Connections for the Fifth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R5 UL to Rack 1	R5-UL-P5 to R1-SS-P8	10 meters
R5 UL to Rack 2	R5-UL-P4 to R2-SS-P8	10 meters
R5 UL to Rack 3	R5-UL-P6 to R3-SS-P12	5 meters
R5 UL to Rack 4	R5-UL-P31 to R4-SS-P7 R5-UL-P33 to R4-SS-P12	5 meters
R5 UL within Rack 5	R5-UL-P30 to R5-SS-P7	3 meters
R5 UL to Rack 6	R5-UL-P7 to R6-SS-P7	5 meters
R5 UL to Rack 7	R5-UL-P32 to R7-SS-P7	5 meters
R5 LL to Rack 1	R5-LL-P7 to R1-SS-P18	10 meters
R5 LL to Rack 2	R5-LL-P6 to R2-SS-P18	10 meters
R5 LL to Rack 3	R5-LL-P31 to R3-SS-P16	5 meters
R5 LL to Rack 4	R5-LL-P33 to R4-SS-P19	5 meters
R5 LL within Rack 5	R5-LL-P30 to R5-SS-P15 R5-LL-P32 to R5-SS-P19	3 meters
R5 LL to Rack 6	R5-LL-P4 to R6-SS-P15	5 meters
R5 LL to Rack 7	R5-LL-P5 to R7-SS-P15	5 meters

The following table describes the cable connections for the sixth spine switch (R6-SS) when cabling seven full racks together:

Table 5-26 Leaf Switch Connections for the Sixth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R6 UL to Rack 1	R6-UL-P7 to R1-SS-P16	10 meters
R6 UL to Rack 2	R6-UL-P4 to R2-SS-P16	10 meters
R6 UL to Rack 3	R6-UL-P6 to R3-SS-P19	10 meters
R6 UL to Rack 4	R6-UL-P31 to R4-SS-P15	5 meters
R6 UL to Rack 5	R6-UL-P33 to R5-SS-P18	5 meters
R6 UL within Rack 6	R6-UL-P30 to R6-SS-P16 R6-UL-P32 to R6-SS-P18	3 meters
R6 UL to Rack 7	R6-UL-P5 to R7-SS-P16	5 meters
R6 LL to Rack 1	R6-LL-P5 to R1-SS-P12 R6-LL-P7 to R1-SS-P15	10 meters
R6 LL to Rack 2	R6-LL-P4 to R2-SS-P12	10 meters
R6 LL to Rack 3	R6-LL-P31 to R3-SS-P7	10 meters
R6 LL to Rack 4	R6-LL-P33 to R4-SS-P9	5 meters
R6 LL to Rack 5	R6-LL-P30 to R5-SS-P8	5 meters
R6 LL within Rack 6	R6-LL-P32 to R6-SS-P8	3 meters
R6 LL to Rack 7	R6-LL-P6 to R7-SS-P8	5 meters

The following table describes the cable connections for the seventh spine switch (R7-SS) when cabling seven full racks together:

Table 5-27 Leaf Switch Connections for the Seventh Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R7 UL to Rack 1	R7-UL-P5 to R1-SS-P9	10 meters
R7 UL to Rack 2	R7-UL-P7 to R2-SS-P7	10 meters
R7 UL to Rack 3	R7-UL-P4 to R3-SS-P9	10 meters
R7 UL to Rack 4	R7-UL-P6 to R4-SS-P8	10 meters
R7 UL to Rack 5	R7-UL-P31 to R5-SS-P9	5 meters
R7 UL to Rack 6	R7-UL-P33 to R6-SS-P9	5 meters
R7 UL within Rack 7	R7-UL-P30 to R7-SS-P9 R7-UL-P32 to R7-SS-P12	3 meters
R7 LL to Rack 1	R7-LL-P5 to R1-SS-P19	10 meters
R7 LL to Rack 2	R7-LL-P7 to R2-SS-P19	10 meters
R7 LL to Rack 3	R7-LL-P4 to R3-SS-P15	10 meters
R7 LL to Rack 4	R7-LL-P6 to R4-SS-P18	10 meters
R7 LL to Rack 5	R7-LL-P31 to R5-SS-P16	5 meters
R7 LL to Rack 6	R7-LL-P33 to R6-SS-P12	5 meters
R7 LL within Rack 7	R7-LL-P30 to R7-SS-P18 R7-LL-P32 to R7-SS-P19	3 meters

5.8 Eight-Rack Cabling for X8M Racks

This section provides the cabling tables to connect eight (8) X8M racks together, all of which use RoCE Network Fabric.

 **Note:**

The following conventions were used in the cabling notation for connecting multiple racks together.

- The spine switch (also referred to as SS) is in U1 for all racks.
- The leaf switches are referred to as Lower Leaf (LL) and Upper Leaf (UL).
- The leaf switches are located as follows:
 - At rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X8M-2 or Storage Expansion Rack X8M-2
 - At rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X8M-8
- The cable lengths shown in the following lists assume that the racks are adjacent to each other, the cables are routed through a raised floor, and there are no obstacles in the routing between the racks. If the racks are not adjacent, or use overhead cabling trays, then they may require longer cable lengths. Cable lengths up to 100 meters are supported.
- Only optical cables (with additional transceivers) are supported for lengths greater than 5 meters.
- For X8M-8 systems with three database servers and 11 storage servers only, you must adjust the following multi-rack cabling information. On such systems only, port 30 on the leaf switches is connected to a database server and is not used as an inter-switch link. Consequently, in the following tables, any connection to port number 30 on any leaf switch must instead connect to port number 34 on the same leaf switch. For example, R1-UL-P30 must change to R1-UL-P34, R1-LL-P30 must change to R1-LL-P34, and so on. These changes only apply to X8M-8 systems with three database servers and 11 storage servers.

The following diagrams show the cable connections for the 8 spine switches when cabling eight racks together:

Rack 1 Spine (R1SS)

1	3	5 R1 UL P5	7 R8 LL P5	9 R7 UL P5	11 R3 UL P5	13 R1 LL P5	15 R8 UL P5	17 R4 LL P5	19 R7 LL P5	21	23	25	27	29	31	33	35
2	4	6 R2 UL P5	8 R5 UL P5	10 R4 UL P7	12 R6 LL P5	14 R2 LL P5	16 R6 UL P7	18 R5 LL P7	20 R3 LL P4	22	24	26	28	30	32	34	36

Rack 2 Spine (R2SS)

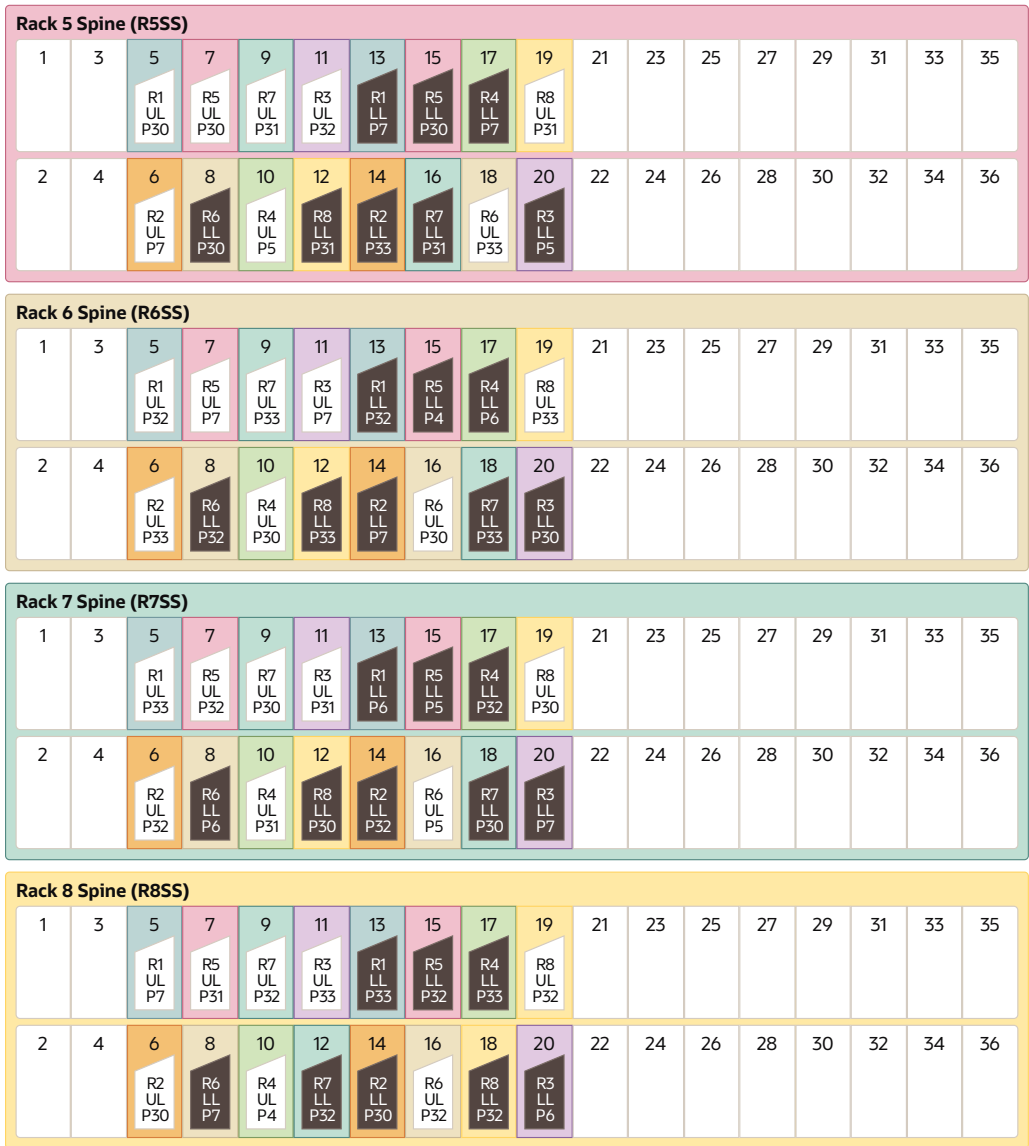
1	3	5 R1 UL P31	7 R7 UL P7	9 R8 LL P7	11 R3 UL P6	13 R1 LL P31	15 R8 UL P7	17 R4 LL P4	19 R7 LL P7	21	23	25	27	29	31	33	35
2	4	6 R2 UL P31	8 R5 UL P4	10 R4 UL P6	12 R6 LL P4	14 R2 LL P31	16 R6 UL P4	18 R5 LL P6	20 R3 LL P31	22	24	26	28	30	32	34	36

Rack 3 Spine (R3SS)

1	3	5 R1 UL P6	7 R6 LL P31	9 R7 UL P4	11 R3 UL P30	13 R1 LL P4	15 R7 LL P4	17 R4 LL P31	19 R6 UL P6	21	23	25	27	29	31	33	35
2	4	6 R2 UL P6	8 R5 UL P6	10 R4 UL P33	12 R8 LL P4	14 R2 LL P6	16 R5 LL P31	18 R8 UL P4	20 R3 LL P32	22	24	26	28	30	32	34	36

Rack 4 Spine (R4SS)

1	3	5 R1 UL P4	7 R8 LL P6	9 R6 LL P33	11 R3 UL P4	13 R1 LL P30	15 R6 UL P31	17 R4 LL P30	19 R5 LL P33	21	23	25	27	29	31	33	35
2	4	6 R2 UL P4	8 R7 UL P6	10 R4 UL P32	12 R5 UL P33	14 R2 LL P4	16 R8 UL P6	18 R7 LL P6	20 R3 LL P33	22	24	26	28	30	32	34	36



The following table describes the cable connections for the first spine switch (R1-SS) when cabling eight racks together:

Table 5-28 Leaf Switch Connections for the First Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R1 UL within Rack 1	R1-UL-P5 to R1-SS-P5	3 meters
R1 UL to Rack 2	R1-UL-P31 to R2-SS-P5	5 meters
R1 UL to Rack 3	R1-UL-P6 to R3-SS-P5	5 meters
R1 UL to Rack 4	R1-UL-P4 to R4-SS-P5	10 meters
R1 UL to Rack 5	R1-UL-P30 to R5-SS-P5	10 meters
R1 UL to Rack 6	R1-UL-P32 to R6-SS-P5	10 meters
R1 UL to Rack 7	R1-UL-P33 to R7-SS-P5	10 meters
R1 UL to Rack 8	R1-UL-P7 to R8-SS-P5	10 meters

Table 5-28 (Cont.) Leaf Switch Connections for the First Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R1 LL within Rack 1	R1-LL-P5 to R1-SS-P13	3 meters
R1 LL to Rack 2	R1-LL-P31 to R2-SS-P13	5 meters
R1 LL to Rack 3	R1-LL-P4 to R3-SS-P13	5 meters
R1 LL to Rack 4	R1-LL-P30 to R4-SS-P13	10 meters
R1 LL to Rack 5	R1-LL-P7 to R5-SS-P13	10 meters
R1 LL to Rack 6	R1-LL-P32 to R6-SS-P13	10 meters
R1 LL to Rack 7	R1-LL-P6 to R7-SS-P13	10 meters
R1 LL to Rack 8	R1-LL-P33 to R8-SS-P13	10 meters

The following table describes the cable connections for the second spine switch (R2-SS) when cabling eight full racks together:

Table 5-29 Leaf Switch Connections for the Second Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R2 UL to Rack 1	R2-UL-P5 to R1-SS-P6	5 meters
R2 UL within Rack 2	R2-UL-P31 to R2-SS-P6	3 meters
R2 UL to Rack 3	R2-UL-P6 to R3-SS-P6	5 meters
R2 UL to Rack 4	R2-UL-P4 to R4-SS-P6	5 meters
R2 UL to Rack 5	R2-UL-P7 to R5-SS-P6	10 meters
R2 UL to Rack 6	R2-UL-P33 to R6-SS-P6	10 meters
R2 UL to Rack 7	R2-UL-P32 to R7-SS-P6	10 meters
R2 UL to Rack 8	R2-UL-P30 to R8-SS-P6	10 meters
R2 LL to Rack 1	R2-LL-P5 to R1-SS-P14	5 meters
R2 LL within Rack 2	R2-LL-P31 to R2-SS-P14	3 meters
R2 LL to Rack 3	R2-LL-P6 to R3-SS-P14	5 meters
R2 LL to Rack 4	R2-LL-P4 to R4-SS-P14	5 meters
R2 LL to Rack 5	R2-LL-P33 to R5-SS-P14	10 meters
R2 LL to Rack 6	R2-LL-P7 to R6-SS-P14	10 meters
R2 LL to Rack 7	R2-LL-P32 to R7-SS-P14	10 meters
R2 LL to Rack 8	R2-LL-P30 to R8-SS-P14	10 meters

The following table describes the cable connections for the third spine switch (R3-SS) when cabling eight full racks together:

Table 5-30 Leaf Switch Connections for the Third Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 1	R3-UL-P5 to R1-SS-P11	5 meters

Table 5-30 (Cont.) Leaf Switch Connections for the Third Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R3 UL to Rack 2	R3-UL-P6 to R2-SS-P11	5 meters
R3 UL within Rack 3	R3-UL-P30 to R3-SS-P11	3 meters
R3 UL to Rack 4	R3-UL-P4 to R4-SS-P11	5 meters
R3 UL to Rack 5	R3-UL-P32 to R5-SS-P11	5 meters
R3 UL to Rack 6	R3-UL-P7 to R6-SS-P11	10 meters
R3 UL to Rack 7	R3-UL-P31 to R7-SS-P11	10 meters
R3 UL to Rack 8	R3-UL-P33 to R8-SS-P11	10 meters
R3 LL to Rack 1	R3-LL-P4 to R1-SS-P20	5 meters
R3 LL to Rack 2	R3-LL-P31 to R2-SS-P20	5 meters
R3 LL within Rack 3	R3-LL-P32 to R3-SS-P20	3 meters
R3 LL to Rack 4	R3-LL-P33 to R4-SS-P20	5 meters
R3 LL to Rack 5	R3-LL-P5 to R5-SS-P20	5 meters
R3 LL to Rack 6	R3-LL-P30 to R6-SS-P20	10 meters
R3 LL to Rack 7	R3-LL-P7 to R7-SS-P20	10 meters
R3 LL to Rack 8	R3-LL-P6 to R8-SS-P20	10 meters

The following table describes the cable connections for the fourth spine switch (R4-SS) when cabling eight full racks together:

Table 5-31 Leaf Switch Connections for the Fourth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R4 UL to Rack 1	R4-UL-P7 to R1-SS-P10	10 meters
R4 UL to Rack 2	R4-UL-P6 to R2-SS-P10	5 meters
R4 UL to Rack 3	R4-UL-P33 to R3-SS-P10	5 meters
R4 UL within Rack 4	R4-UL-P32 to R4-SS-P10	3 meters
R4 UL to Rack 5	R4-UL-P5 to R5-SS-P10	5 meters
R4 UL to Rack 6	R4-UL-P30 to R6-SS-P10	5 meters
R4 UL to Rack 7	R4-UL-P31 to R7-SS-P10	10 meters
R4 UL to Rack 8	R4-UL-P4 to R8-SS-P10	10 meters
R4 LL to Rack 1	R4-LL-P5 to R1-SS-P17	10 meters
R4 LL to Rack 2	R4-LL-P4 to R2-SS-P17	5 meters
R4 LL to Rack 3	R4-LL-P31 to R3-SS-P17	5 meters
R4 LL within Rack 4	R4-LL-P30 to R4-SS-P17	3 meters
R4 LL to Rack 5	R4-LL-P7 to R5-SS-P17	5 meters
R4 LL to Rack 6	R4-LL-P6 to R6-SS-P17	5 meters
R4 LL to Rack 7	R4-LL-P32 to R7-SS-P17	10 meters

Table 5-31 (Cont.) Leaf Switch Connections for the Fourth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R4 LL to Rack 8	R4-LL-P33 to R8-SS-P17	10 meters

The following table describes the cable connections for the fifth spine switch (R5-SS) when cabling eight full racks together:

Table 5-32 Leaf Switch Connections for the Fifth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R5 UL to Rack 1	R5-UL-P5 to R1-SS-P8	10 meters
R5 UL to Rack 2	R5-UL-P4 to R2-SS-P8	10 meters
R5 UL to Rack 3	R5-UL-P6 to R3-SS-P8	5 meters
R5 UL to Rack 4	R5-UL-P33 to R4-SS-P12	5 meters
R5 UL within Rack 5	R5-UL-P30 to R5-SS-P7	3 meters
R5 UL to Rack 6	R5-UL-P7 to R6-SS-P7	5 meters
R5 UL to Rack 7	R5-UL-P32 to R7-SS-P7	5 meters
R5 UL to Rack 8	R5-UL-P31 to R8-SS-P7	10 meters
R5 LL to Rack 1	R5-LL-P7 to R1-SS-P18	10 meters
R5 LL to Rack 2	R5-LL-P6 to R2-SS-P18	10 meters
R5 LL to Rack 3	R5-LL-P31 to R3-SS-P16	5 meters
R5 LL to Rack 4	R5-LL-P33 to R4-SS-P19	5 meters
R5 LL within Rack 5	R5-LL-P30 to R5-SS-P15	3 meters
R5 LL to Rack 6	R5-LL-P4 to R6-SS-P15	5 meters
R5 LL to Rack 7	R5-LL-P5 to R7-SS-P15	5 meters
R5 LL to Rack 8	R5-LL-P32 to R8-SS-P15	10 meters

The following table describes the cable connections for the sixth spine switch (R6-SS) when cabling eight full racks together:

Table 5-33 Leaf Switch Connections for the Sixth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R6 UL to Rack 1	R6-UL-P7 to R1-SS-P16	10 meters
R6 UL to Rack 2	R6-UL-P4 to R2-SS-P16	10 meters
R6 UL to Rack 3	R6-UL-P6 to R3-SS-P19	10 meters
R6 UL to Rack 4	R6-UL-P31 to R4-SS-P15	5 meters
R6 UL to Rack 5	R6-UL-P33 to R5-SS-P18	5 meters
R6 UL within Rack 6	R6-UL-P30 to R6-SS-P16	3 meters
R6 UL to Rack 7	R6-UL-P5 to R7-SS-P16	5 meters
R6 UL to Rack 8	R6-UL-P32 to R8-SS-P16	5 meters

Table 5-33 (Cont.) Leaf Switch Connections for the Sixth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R6 LL to Rack 1	R6-LL-P5 to R1-SS-P12	10 meters
R6 LL to Rack 2	R6-UL-P4 to R2-SS-P12	10 meters
R6 LL to Rack 3	R6-LL-P31 to R3-SS-P7	10 meters
R6 LL to Rack 4	R6-LL-P33 to R4-SS-P9	5 meters
R6 LL to Rack 5	R6-LL-P30 to R5-SS-P8	5 meters
R6 LL within Rack 6	R6-LL-P32 to R6-SS-P8	3 meters
R6 LL to Rack 7	R6-LL-P6 to R7-SS-P8	5 meters
R6 LL to Rack 8	R6-LL-P7 to R8-SS-P8	5 meters

The following table describes the cable connections for the seventh spine switch (R7-SS) when cabling eight full racks together:

Table 5-34 Leaf Switch Connections for the Seventh Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R7 UL to Rack 1	R7-UL-P5 to R1-SS-P9	10 meters
R7 UL to Rack 2	R7-UL-P7 to R2-SS-P7	10 meters
R7 UL to Rack 3	R7-UL-P4 to R3-SS-P9	10 meters
R7 UL to Rack 4	R7-UL-P6 to R4-SS-P8	10 meters
R7 UL to Rack 5	R7-UL-P31 to R5-SS-P9	5 meters
R7 UL to Rack 6	R7-UL-P33 to R6-SS-P9	5 meters
R7 UL within Rack 7	R7-UL-P30 to R7-SS-P9	3 meters
R7 UL to Rack 8	R7-UL-P32 to R8-SS-P9	5 meters
R7 LL to Rack 1	R7-LL-P5 to R1-SS-P19	10 meters
R7 LL to Rack 2	R7-LL-P7 to R2-SS-P19	10 meters
R7 LL to Rack 3	R7-LL-P4 to R3-SS-P15	10 meters
R7 LL to Rack 4	R7-LL-P6 to R4-SS-P18	10 meters
R7 LL to Rack 5	R7-LL-P31 to R5-SS-P16	5 meters
R7 LL to Rack 6	R7-LL-P33 to R6-SS-P18	5 meters
R7 LL within Rack 7	R7-LL-P30 to R7-SS-P18	3 meters
R7 LL to Rack 8	R7-LL-P32 to R8-SS-P12	5 meters

The following table describes the cable connections for the eighth spine switch (R8-SS) when cabling eight full racks together:

Table 5-35 Leaf Switch Connections for the Eighth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R8 UL to Rack 1	R8-UL-P5 to R1-SS-P15	10 meters
R8 UL to Rack 2	R8-UL-P7 to R2-SS-P15	10 meters
R8 UL to Rack 3	R8-UL-P4 to R3-SS-P18	10 meters
R8 UL to Rack 4	R8-UL-P6 to R4-SS-P16	10 meters
R8 UL to Rack 5	R8-UL-P31 to R5-SS-P19	10 meters
R8 UL to Rack 6	R8-UL-P33 to R6-SS-P19	5 meters
R8 UL to Rack 7	R8-UL-P30 to R7-SS-P19	5 meters
R8 UL within Rack 8	R8-UL-P32 to R8-SS-P19	3 meters
R8 LL to Rack 1	R8-LL-P5 to R1-SS-P7	10 meters
R8 LL to Rack 2	R8-LL-P7 to R2-SS-P9	10 meters
R8 LL to Rack 3	R8-LL-P4 to R3-SS-P12	10 meters
R8 LL to Rack 4	R8-LL-P6 to R4-SS-P7	10 meters
R8 LL to Rack 5	R8-LL-P31 to R5-SS-P12	10 meters
R8 LL to Rack 6	R8-LL-P33 to R6-SS-P12	5 meters
R8 LL to Rack 7	R8-LL-P30 to R7-SS-P12	5 meters
R8 LL within Rack 8	R8-LL-P32 to R8-SS-P18	3 meters

6

Multi-Rack Cabling Tables for Oracle Exadata Rack Models with InfiniBand Network Fabric (X2 to X8)

This section contains multi-rack cabling tables for Oracle Exadata Rack models that use InfiniBand Network Fabric. This includes Oracle Exadata Rack models from X2 to X8.

- [Understanding Multi-Rack Cabling for Racks with InfiniBand Network Fabric](#)
Up to eight racks can be cabled together without external RDMA Network Fabric switches.
- [Two-Rack Cabling with InfiniBand Network Fabric](#)
Review this information before cabling two racks together with InfiniBand Network Fabric.
- [Three-Rack Cabling with InfiniBand Network Fabric](#)
- [Four-Rack Cabling with InfiniBand Network Fabric](#)
- [Five-Rack Cabling with InfiniBand Network Fabric](#)
- [Six-Rack Cabling with InfiniBand Network Fabric](#)
- [Seven-Rack Cabling with InfiniBand Network Fabric](#)
- [Eight-Rack Cabling with InfiniBand Network Fabric](#)

6.1 Understanding Multi-Rack Cabling for Racks with InfiniBand Network Fabric

Up to eight racks can be cabled together without external RDMA Network Fabric switches.

Three RDMA Network Fabric switches are needed when cabling racks together. If any rack does not include a third switch, then order the expansion switch kit. These switches attach to standard RDMA Network Fabric connectors at the end of the cables. The procedures in this section assume the racks are adjacent to each other. If they are not, then longer cables may be required for the connections.

 **Note:**

Only for InfiniBand Network Fabric (X8 and earlier).

- For Eighth or Quarter Racks, which are the smallest Elastic Configurations, follow the instructions in "[Cabling Oracle Exadata Quarter Racks and Oracle Exadata Eighth Racks with InfiniBand Network Fabric](#)" for direct connection without spine switch.
- For other racks (Half Rack, Full Rack, Elastic Configurations larger than Eighth or Quarter Rack) install a spine switch and follow the standard *N*-rack cabling using fat tree topology for two or more interconnected racks, for example "[Two-Rack Cabling with InfiniBand Network Fabric](#)."

The switch at rack unit 1 (U1) is referred to as the spine switch.

The following switches are referred to as leaf switches:

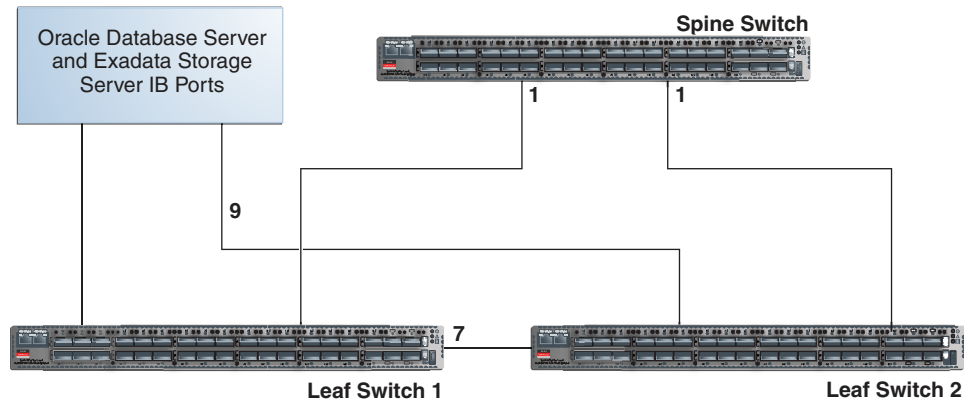
- The switches at rack unit 20 (U20) and rack unit 24 (U24) in Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks
- The switches at rack unit 21 (U21) and rack unit 23 (U23) in Exadata Database Machine X2-8 and later.
- The switches at rack unit 20 (U20) and 22 (U22) in Exadata Database Machine X3-2 and later racks or Storage Expansion Rack X3-2 and later.

 **Note:**

Oracle Exadata Database Machine X4-2 and later racks or Oracle Exadata Database Machine X3-8 Full Racks with Exadata Storage Server X4-2L Servers do not include spare cables or a third Sun Datacenter InfiniBand Switch 36 switch. To extend Oracle Exadata Database Machine X4-2 and later racks or Oracle Exadata Database Machine X3-8 Full Racks with Exadata Storage Server X4-2L Servers, you must order cables and a Sun Datacenter InfiniBand Switch 36 switch.

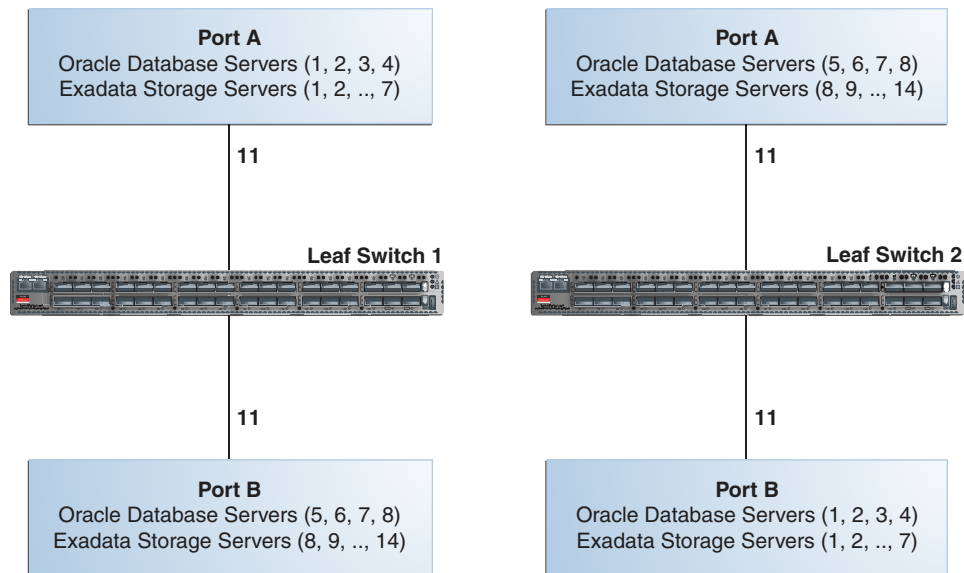
In a single rack, the two leaf switches are interconnected using seven connections. In addition, each leaf switch has one connection to the spine switch. The leaf switches connect to the spine switch as shown in the following graphic:

Figure 6-1 Connections Between Spine Switch and Leaf Switches in a Single Rack



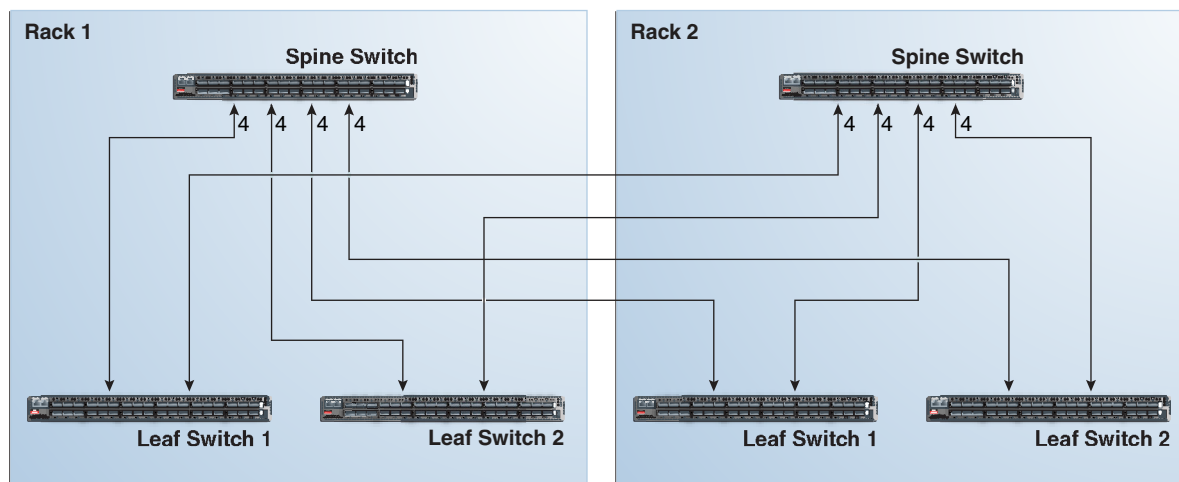
The Oracle Database servers and Exadata Storage Servers connect to the leaf switches as shown in the following graphic:

Figure 6-2 Connections Between Database Servers and Storage Servers and Leaf Switches



When connecting up to eight racks together, remove the seven existing inter-switch connections between each leaf switch, as well as the two connections between the leaf switches and the spine switch. From each leaf switch, distribute eight connections over the spine switches in all racks. In multi-rack environments, the leaf switches inside a rack are no longer directly interconnected, as shown in the following graphic:

Figure 6-3 Connections Between Spine Switches and Leaf Switches Across Two Racks



As shown in the preceding graphic, each leaf switch in rack 1 connects to the following switches:

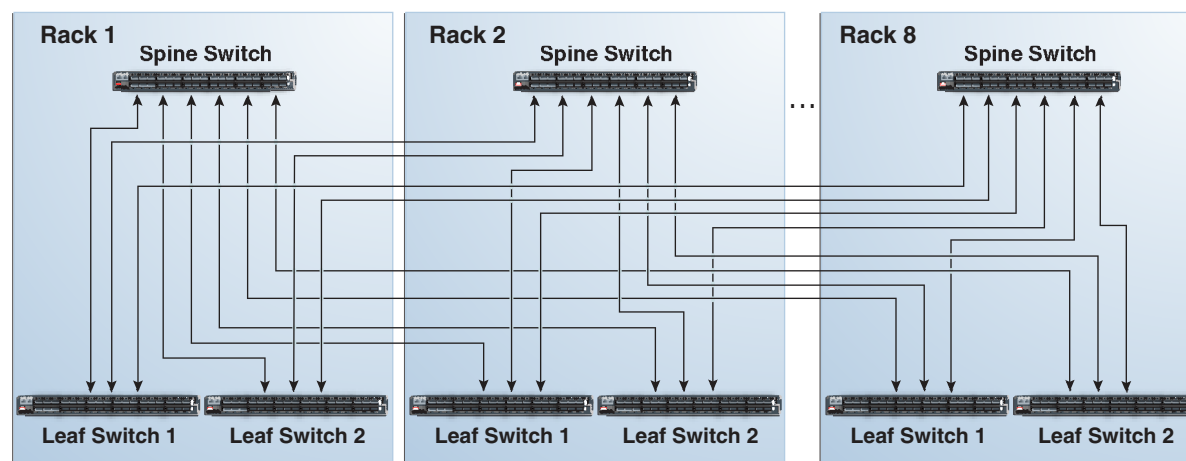
- Four connections to its internal spine switch
- Four connections to the spine switch in rack 2

The spine switch in rack 1 connects to the following switches:

- Eight connections to both internal leaf switches
- Eight connections to both leaf switches in rack 2

As the number of racks increases from two to eight, the pattern continues as shown in the following graphic:

Figure 6-4 Connections Between Spine Switches and Leaf Switches for up to 8 Racks



As shown in the preceding graphic, each leaf switch has eight inter-switch connections distributed over all spine switches. Each spine switch has 16 inter-switch connections distributed over all leaf switches. The leaf switches are not directly interconnected with

other leaf switches, and the spine switches are not directly interconnected with the other spine switches.

**Note:**

Cable lengths are specified for racks 1 through 8.

- [Preparing for Multi-Rack Cabling with InfiniBand Network Fabric](#)
- [Cabling Oracle Exadata Quarter Racks and Oracle Exadata Eighth Racks with InfiniBand Network Fabric](#)

6.1.1 Preparing for Multi-Rack Cabling with InfiniBand Network Fabric

Racks can be added together to increase system capacity and performance. When cabling racks together, note the following:

- The cable lengths shown in this document assume the racks are adjacent to each other. If the racks are not adjacent, or there are obstacles in the raised floor, or if you use overhead cabling, then longer cables may be required. For optical cables, the maximum supported cable length is 100 meters. For copper cables, the maximum supported cable length is 5 meters.
- Oracle recommends that the names for the servers include the rack unit number. This helps identify the server during diagnostics.
- When completing Oracle Exadata Deployment Assistant for the additional rack, you are prompted for SCAN addresses. However, these SCAN addresses are not used because the SCAN address from the original rack are used. Manually remove the new SCAN addresses from the generated installation files.
- The software owner account names and group names, as well as their identifiers, must match the names and identifiers of the original rack.
- If the additional grid disks are used with existing disk groups, then ensure the grid disk sizes for the new rack are the same as the original rack.
- If the InfiniBand network consists of four or more racks cabled together, then disable the Subnet Manager on the leaf switches.
- Verify the Master Subnet Manager is located on the spine switch.
- Oracle Exadata Database Machine Quarter Racks can be extended as follows:
 - Connect two Oracle Exadata Database Machine Quarter Racks together. At least four of the six ports reserved for external connectivity are open on each leaf switch. The six ports are 5B, 6A, 6B, 7A, 7B, and 12A in each leaf switch. Maintain the existing seven inter-switch links between the leaf switches within each rack. Connect the leaf switches between the racks with two links each, using the ports reserved for external connectivity.
 - Connect one Oracle Exadata Database Machine Quarter Rack with one Oracle Exadata Database Machine Half Rack or one Oracle Exadata Database Machine Full Rack. At least four ports reserved for external connectivity are open on each leaf switch. The spine switch in the Oracle Exadata Database Machine Half Rack or Oracle Exadata Database Machine Full Rack remains as the spine switch. Maintain the existing seven inter-switch links between the leaf switches within each rack.

Connect the leaf switches between the racks with two links each, using the ports reserved for external connectivity.

- Connect one Oracle Exadata Database Machine Quarter Rack with two or more Oracle Exadata Database Machine Half Racks or Oracle Exadata Database Machine Full Racks. The racks are interconnected using a fat-tree topology. Connect each leaf switch in the quarter rack to the spine switch of each half rack or full rack using two links each. If there are more than four racks, then use one link instead of two. The seven inter-switch links between the leaf switches in the quarter rack are removed.

 **Note:**

To connect more than one quarter rack to additional racks, it is necessary to purchase Sun Datacenter InfiniBand Switch 36 switches for the quarter racks.

- If you are extending Oracle Exadata Database Machine X4-2 or later, or Oracle Exadata Database Machine X3-8 Full Rack, or Oracle Exadata Database Machine X2-2 (with X4170 and X4275 servers) half rack, then order the expansion kit that includes a Sun Datacenter InfiniBand Switch 36 switch.

Perform the following tasks before cabling racks together:

1. Determine the number of racks that will be cabled together.
2. Count the spare cables from the kit, and existing inter-switch cables.

 **Note:**

Oracle Exadata Database Machine X4-2 and later racks or Oracle Exadata Database Machine X3-8 Full Racks with Exadata Storage Server X4-2L Servers do not include spare cables or a third Sun Datacenter InfiniBand Switch 36 switch. To extend Oracle Exadata Database Machine X4-2 and later racks or Oracle Exadata Database Machine X3-8 Full Racks with Exadata Storage Server X4-2L Servers, you must order cables and a Sun Datacenter InfiniBand Switch 36 switch.

For Oracle Exadata Racks earlier than Oracle Exadata Database Machine X4-2, no additional InfiniBand cables need to be purchased when connecting up to three Oracle Exadata Database Machine Full Racks. The following table lists the spare cables for the switch:

Cable	Spare	Inter-switch	Total
InfiniBand 2 meter cable	0	7	7
InfiniBand 3 meter cable	6	2	8
InfiniBand 5 meter cable	10	0	10

When connecting four or more racks, it is necessary to purchase additional 10 meter optical QSFP-to-QSFP QDR InfiniBand cables. Two QSFP SR transceivers are needed for each optical cable. The following table lists the number of racks, and cables needed to interconnect the racks.

Number of Racks	3 Meter Cables Needed	5 Meter Cables Needed	10 Meter Cables Needed
2	16	16	0
3	18	30	0
4	16	40	8
5	20	42	18
6	24	46	26
7	28	44	40
8	16	52	60

3. Gather the following tools:
 - Pen
 - Note pad
 - Flashlight
 - Labels for cables and rack
 - Cable ties (Oracle recommends hook and loop straps)
4. Determine a naming method for the rack prefixes. For example, if the original rack has the prefix `dbm01`, then use the prefix `dbm02` for the second rack, the prefix `dbm03` for the third rack, and so on.
5. Verify the racks have unique host names and IP addresses. All servers interconnected in the racks must have unique names and IP addresses.

Server names and IP addresses conventions may differ in the following cases:

- Initial installation of all Oracle Exadata Database Machine Full Racks: System address assignments and host names should be complete.
 - New Oracle Exadata Database Machine Full Racks are added to an existing cluster: The new rack configuration should require unique host names and IP addresses for the new Oracle Exadata Database Machines. The IP addresses on the same subnet cannot conflict with the existing systems.
 - Two existing Oracle Exadata Database Machine Full Racks are clustered together: You can assign host names and IP addresses only if Oracle Exadata Database Machines are already assigned unique host names and IP addresses, or the entire cluster must be reconfigured. The machines must be on the same subnet and not have conflicting IP addresses.
6. Ensure the IP addresses for the new servers are in the same subnet, and do not overlap with the currently-installed servers.
 7. Ensure the firmware on the original switches are at the same level as the new switches using the `nm2version` command. If the firmware is not at the same level, then apply a firmware patch.

6.1.2 Cabling Oracle Exadata Quarter Racks and Oracle Exadata Eighth Racks with InfiniBand Network Fabric

Oracle Exadata Quarter Racks and Oracle Exadata Eighth Racks with InfiniBand Network Fabric can be cabled as follows:

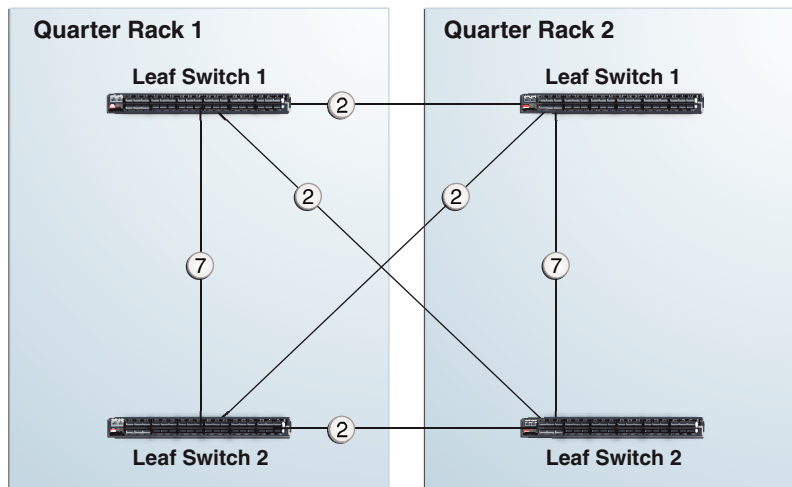
- Oracle Exadata Quarter Rack to Oracle Exadata Quarter Rack
- Oracle Exadata Quarter Rack to Oracle Exadata Half Rack, or multiple Oracle Exadata Half Racks or Oracle Exadata Full Racks
- Oracle Exadata Quarter Rack to Oracle Exadata Full Rack, or multiple Oracle Exadata Full Racks or Oracle Exadata Half Racks
- Oracle Exadata Eighth Rack to Oracle Exadata Eighth Rack
- Oracle Exadata Eighth Rack to Oracle Exadata Half Rack, or multiple Oracle Exadata Half Racks or Oracle Exadata Full Racks
- Oracle Exadata Eighth Rack to Oracle Exadata Full Rack, or multiple Oracle Exadata Full Racks or Oracle Exadata Half Racks

Note:

- To cable additional Oracle Exadata Quarter Racks or Oracle Exadata Eighth Racks, it is necessary to purchase additional Sun Datacenter InfiniBand Switch 36 switches.
- It is assumed that four of the six ports dedicated for external connectivity in Oracle Exadata Half Rack or Oracle Exadata Full Rack are unused and available for switch interconnections with Oracle Exadata Quarter Rack or Oracle Exadata Eighth Rack. The six ports are 5B, 6A, 6B, 7A, 7B, and 12A in each leaf switch.

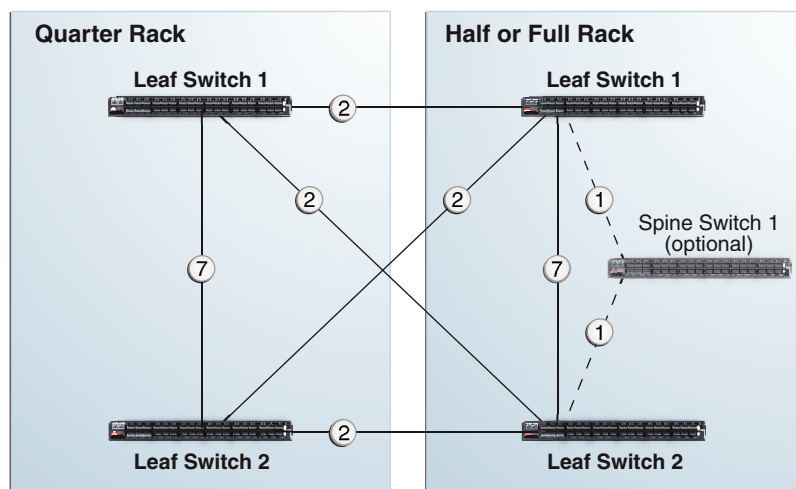
The following graphic shows the cable connections for two Oracle Exadata Quarter Racks. The leaf switches within each rack maintain their existing seven connections. The leaf switches interconnect between the racks with two links each using the ports reserved for external connectivity.

Figure 6-5 Leaf Switch Connections Between 2 Quarter Racks



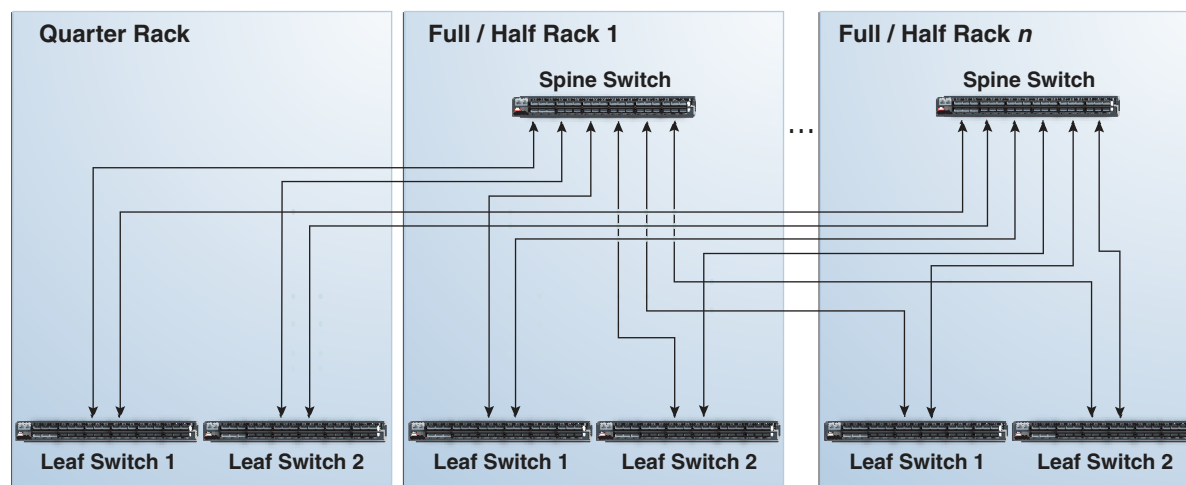
The following graphic shows the cable connections from Oracle Exadata Quarter Rack to Oracle Exadata Half Rack or Oracle Exadata Full Rack. The leaf switches within each rack maintain their existing seven connections. The leaf switches interconnect between the racks with two links each using the ports reserved for external connectivity.

Figure 6-6 Leaf and Spine Switch Connections Between a Quarter Rack and a Half or Full Rack



The following graphic shows the cable connections from Oracle Exadata Quarter Rack to two or more racks. The racks that connect to Oracle Exadata Quarter Rack must be all Oracle Exadata Half Racks or Oracle Exadata Full Racks, interconnected using a fat-tree topology. Each leaf switch in Oracle Exadata Quarter Rack connects to the spine switches in the other half racks or full racks with two links each. If there are more than four racks, then use one link instead of two. The seven inter-switch links between the leaf switches in the quarter rack are removed.

Figure 6-7 Leaf and Spine Switch Connections for a Quarter Rack Connected to One or More Half or Full Racks



6.2 Two-Rack Cabling with InfiniBand Network Fabric

Review this information before cabling two racks together with InfiniBand Network Fabric.

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following lists, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine 8-socket (X8-8, X7-8, X6-8, X5-8, X4-8, X3-8 and X2-8) Full Rack systems, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later racks, Oracle Exadata Database Machine X3-2 Full Racks and Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the following lists assume that the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Cables lengths up to 100 meters are supported.

The following table shows the cable connections for the first spine switch (R1-U1) when cabling two full racks together.

Table 6-1 Leaf Switch Connections for the First Rack in a Two-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A R1-IB3-P8B to R1-U1-P4A R1-IB3-P9A to R1-U1-P5A R1-IB3-P9B to R1-U1-P6A	3 meters
R1 IB3 to Rack 2	R1-IB3-P10A to R2-U1-P7A R1-IB3-P10B to R2-U1-P8A R1-IB3-P11A to R2-U1-P9A R1-IB3-P11B to R2-U1-P10A	5 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B R1-IB2-P8B to R1-U1-P4B R1-IB2-P9A to R1-U1-P5B R1-IB2-P9B to R1-U1-P6B	3 meters
R1 IB2 to Rack 2	R1-IB2-P10A to R2-U1-P7B R1-IB2-P10B to R2-U1-P8B R1-IB2-P11A to R2-U1-P9B R1-IB2-P11B to R2-U1-P10B	5 meters

The following table shows the cable connections for the second spine switch (R2-U1) when cabling two full racks together.

Table 6-2 Leaf Switch Connections for the Second Rack in a Two-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A R2-IB3-P8B to R2-U1-P4A R2-IB3-P9A to R2-U1-P5A R2-IB3-P9B to R2-U1-P6A	3 meters
R2 IB3 to Rack 1	R2-IB3-P10A to R1-U1-P7A R2-IB3-P10B to R1-U1-P8A R2-IB3-P11A to R1-U1-P9A R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B R2-IB2-P8B to R2-U1-P4B R2-IB2-P9A to R2-U1-P5B R2-IB2-P9B to R2-U1-P6B	3 meters

Table 6-2 (Cont.) Leaf Switch Connections for the Second Rack in a Two-Rack System

Leaf Switch	Connection	Cable Length
R2 IB2 to Rack 1	R2-IB2-P10A to R1-U1-P7B R2-IB2-P10B to R1-U1-P8B R2-IB2-P11A to R1-U1-P9B R2-IB2-P11B to R1-U1-P10B	5 meters

Related Topics

- [Cabling Two Racks Together](#)
Choose from the available methods based on your system specifications and operational requirements.

6.3 Three-Rack Cabling with InfiniBand Network Fabric

Note the following when cabling three racks together with InfiniBand Network Fabric:

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following tables, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X2-8 and later, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later racks, or Oracle Exadata Database Machine X3-2 Full Racks or Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the tables assume the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Up to 100 meters is supported.
- Only optical cables are supported for lengths greater than 5 meters.

[Table 6-3](#) shows the cable connections for the first spine switch (R1-U1) when cabling three racks together:

Table 6-3 Leaf Switch Connections for the First Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A R1-IB3-P8B to R1-U1-P4A R1-IB3-P9A to R1-U1-P5A	3 meters

Table 6-3 (Cont.) Leaf Switch Connections for the First Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 to Rack 2	R1-IB3-P9B to R2-U1-P6A R1-IB3-P10A to R2-U1-P7A R1-IB3-P10B to R2-U1-P8A	5 meters
R1 IB3 to Rack 3	R1-IB3-P11A to R3-U1-P9A R1-IB3-P11B to R3-U1-P10A	5 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B R1-IB2-P8B to R1-U1-P4B R1-IB2-P9A to R1-U1-P5B	3 meters
R1 IB2 to Rack 2	R1-IB2-P9B to R2-U1-P6B R1-IB2-P10A to R2-U1-P7B R1-IB2-P10B to R2-U1-P8B	5 meters
R1 IB2 to Rack 3	R1-IB2-P11A to R3-U1-P9B R1-IB2-P11B to R3-U1-P10B	5 meters

Table 6-4 shows the cable connections for the second spine switch (R2-U1) when cabling three racks together:

Table 6-4 Leaf Switch Connections for the Second Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A R2-IB3-P8B to R2-U1-P4A R2-IB3-P9A to R2-U1-P5A	3 meters
R2 IB3 to Rack 1	R2-IB3-P11A to R1-U1-P9A R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB3 to Rack 3	R2-IB3-P9B to R3-U1-P6A R2-IB3-P10A to R3-U1-P7A R2-IB3-P10B to R3-U1-P8A	5 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B R2-IB2-P8B to R2-U1-P4B R2-IB2-P9A to R2-U1-P5B	3 meters
R2 IB2 to Rack 1	R2-IB2-P11A to R1-U1-P9B R2-IB2-P11B to R1-U1-P10B	5 meters
R2 IB2 to Rack 3	R2-IB2-P9B to R3-U1-P6B R2-IB2-P10A to R3-U1-P7B R2-IB2-P10B to R3-U1-P8B	5 meters

Table 6-5 shows the cable connections for the third spine switch (R3-U1) when cabling three full racks together:

Table 6-5 Leaf Switch Connections for the Third Rack in a Three-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 within Rack 3	R3-IB3-P8A to R3-U1-P3A R3-IB3-P8B to R3-U1-P4A R3-IB3-P9A to R3-U1-P5A	3 meters
R3 IB3 to Rack 1	R3-IB3-P9B to R1-U1-P6A R3-IB3-P10A to R1-U1-P7A R3-IB3-P10B to R1-U1-P8A	5 meters
R3 IB3 to Rack 2	R3-IB3-P11A to R2-U1-P9A R3-IB3-P11B to R2-U1-P10A	5 meters
R3 IB2 within Rack 3	R3-IB2-P8A to R3-U1-P3B R3-IB2-P8B to R3-U1-P4B R3-IB2-P9A to R3-U1-P5B	3 meters
R3 IB2 to Rack 1	R3-IB2-P9B to R1-U1-P6B R3-IB2-P10A to R1-U1-P7B R3-IB2-P10B to R1-U1-P8B	5 meters
R3 IB2 to Rack 2	R3-IB2-P11A to R2-U1-P9B R3-IB2-P11B to R2-U1-P10B	5 meters

6.4 Four-Rack Cabling with InfiniBand Network Fabric

Note the following when cabling four racks together with InfiniBand Network Fabric:

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following tables, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X2-8 and later, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later racks, or Oracle Exadata Database Machine X3-2 Full Racks or Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the tables assume the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Up to 100 meters is supported.
- Only optical cables are supported for lengths greater than 5 meters.

Table 6-6 shows the cable connections for the first spine switch (R1-U1) when cabling four racks together:

Table 6-6 Leaf Switch Connections for the First Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A R1-IB3-P8B to R1-U1-P4A	3 meters
R1 IB3 to Rack 2	R1-IB3-P9A to R2-U1-P5A R1-IB3-P9B to R2-U1-P6A	5 meters
R1 IB3 to Rack 3	R1-IB3-P10A to R3-U1-P7A R1-IB3-P10B to R3-U1-P8A	5 meters
R1 IB3 to Rack 4	R1-IB3-P11A to R4-U1-P9A R1-IB3-P11B to R4-U1-P10A	10 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B R1-IB2-P8B to R1-U1-P4B	3 meters
R1 IB2 to Rack 2	R1-IB2-P9A to R2-U1-P5B R1-IB2-P9B to R2-U1-P6B	5 meters
R1 IB2 to Rack 3	R1-IB2-P10A to R3-U1-P7B R1-IB2-P10B to R3-U1-P8B	5 meters
R1 IB2 to Rack 4	R1-IB2-P11A to R4-U1-P9B R1-IB2-P11B to R4-U1-P10B	10 meters

Table 6-7 shows the cable connections for the second spine switch (R2-U1) when cabling four full racks together:

Table 6-7 Leaf Switch Connections for the Second Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A R2-IB3-P8B to R2-U1-P4A	3 meters
R2 IB3 to Rack 1	R2-IB3-P11A to R1-U1-P9A R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB3 to Rack 3	R2-IB3-P9A to R3-U1-P5A R2-IB3-P9B to R3-U1-P6A	5 meters
R2 IB3 to Rack 4	R2-IB3-P10A to R4-U1-P7A R2-IB3-P10B to R4-U1-P8A	5 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B R2-IB2-P8B to R2-U1-P4B	3 meters
R2 IB2 to Rack 1	R2-IB2-P11A to R1-U1-P9B R2-IB2-P11B to R1-U1-P10B	5 meters
R2 IB2 to Rack 3	R2-IB2-P9A to R3-U1-P5B R2-IB2-P9B to R3-U1-P6B	5 meters
R2 IB2 to Rack 4	R2-IB2-P10A to R4-U1-P7B R2-IB2-P10B to R4-U1-P8B	5 meters

Table 6-8 shows the cable connections for the third spine switch (R3-U1) when cabling four full racks together:

Table 6-8 Leaf Switch Connections for the Third Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 within Rack 3	R3-IB3-P8A to R3-U1-P3A R3-IB3-P8B to R3-U1-P4A	3 meters
R3 IB3 to Rack 1	R3-IB3-P10A to R1-U1-P7A R3-IB3-P10B to R1-U1-P8A	5 meters
R3 IB3 to Rack 2	R3-IB3-P11A to R2-U1-P9A R3-IB3-P11B to R2-U1-P10A	5 meters
R3 IB3 to Rack 4	R3-IB3-P9A to R4-U1-P5A R3-IB3-P9B to R4-U1-P6A	5 meters
R3 IB2 within Rack 3	R3-IB2-P8A to R3-U1-P3B R3-IB2-P8B to R3-U1-P4B	3 meters
R3 IB2 to Rack 1	R3-IB2-P10A to R1-U1-P7B R3-IB2-P10B to R1-U1-P8B	5 meters
R3 IB2 to Rack 2	R3-IB2-P11A to R2-U1-P9B R3-IB2-P11B to R2-U1-P10B	5 meters
R3 IB2 to Rack 4	R3-IB2-P9A to R4-U1-P5B R3-IB2-P9B to R4-U1-P6B	5 meters

Table 6-9 shows the cable connections for the fourth spine switch (R4-U1) when cabling four full racks together:

Table 6-9 Leaf Switch Connections for the Fourth Rack in a Four-Rack System

Leaf Switch	Connection	Cable Length
R4 IB3 within Rack 4	R4-IB3-P8A to R4-U1-P3A R4-IB3-P8B to R4-U1-P4A	3 meters
R4 IB3 to Rack 1	R4-IB3-P9A to R1-U1-P5A R4-IB3-P9B to R1-U1-P6A	10 meters
R4 IB3 to Rack 2	R4-IB3-P10A to R2-U1-P7A R4-IB3-P10B to R2-U1-P8A	5 meters
R4 IB3 to Rack 3	R4-IB3-P11A to R3-U1-P9A R4-IB3-P11B to R3-U1-P10A	5 meters
R4 IB2 within Rack 4	R4-IB2-P8A to R4-U1-P3B R4-IB2-P8B to R4-U1-P4B	3 meters
R4 IB2 to Rack 1	R4-IB2-P9A to R1-U1-P5B R4-IB2-P9B to R1-U1-P6B	10 meters
R4 IB2 to Rack 2	R4-IB2-P10A to R2-U1-P7B R4-IB2-P10B to R2-U1-P8B	5 meters
R4 IB2 to Rack 3	R4-IB2-P11A to R3-U1-P9B R4-IB2-P11B to R3-U1-P10B	5 meters

6.5 Five-Rack Cabling with InfiniBand Network Fabric

Note the following when cabling five racks together with InfiniBand Network Fabric:

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following tables, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X2-8 and later racks, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later racks, or Oracle Exadata Database Machine X3-2 Full Racks and Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the tables assume the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Up to 100 meters is supported.
- Only optical cables are supported for lengths greater than 5 meters.

Table 6-10 shows the cable connections for the first spine switch (R1-U1) when cabling five racks together:

Table 6-10 Leaf Switch Connections for the First Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A R1-IB3-P8B to R1-U1-P4A	3 meters
R1 IB3 to Rack 2	R1-IB3-P9A to R2-U1-P5A R1-IB3-P9B to R2-U1-P6A	5 meters
R1 IB3 to Rack 3	R1-IB3-P10A to R3-U1-P7A R1-IB3-P10B to R3-U1-P8A	5 meters
R1 IB3 to Rack 4	R1-IB3-P11A to R4-U1-P9A	10 meters
R1 IB3 to Rack 5	R1-IB3-P11B to R5-U1-P10A	10 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B R1-IB2-P8B to R1-U1-P4B	3 meters
R1 IB2 to Rack 2	R1-IB2-P9A to R2-U1-P5B R1-IB2-P9B to R2-U1-P6B	5 meters
R1 IB2 to Rack 3	R1-IB2-P10A to R3-U1-P7B R1-IB2-P10B to R3-U1-P8B	5 meters
R1 IB2 to Rack 4	R1-IB2-P11A to R4-U1-P9B	10 meters
R1 IB2 to Rack 5	R1-IB2-P11B to R5-U1-P10B	10 meters

Table 6-11 shows the cable connections for the second spine switch (R2-U1) when cabling five full racks together:

Table 6-11 Leaf Switch Connections for the Second Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A R2-IB3-P8B to R2-U1-P4A	3 meters
R2 IB3 to Rack 1	R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB3 to Rack 3	R2-IB3-P9A to R3-U1-P5A R2-IB3-P9B to R3-U1-P6A	5 meters
R2 IB3 to Rack 4	R2-IB3-P10A to R4-U1-P7A R2-IB3-P10B to R4-U1-P8A	5 meters
R2 IB3 to Rack 5	R2-IB3-P11A to R5-U1-P9A	10 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B R2-IB2-P8B to R2-U1-P4B	3 meters
R2 IB2 to Rack 1	R2-IB2-P11B to R1-U1-P10B	5 meters
R2 IB2 to Rack 3	R2-IB2-P9A to R3-U1-P5B R2-IB2-P9B to R3-U1-P6B	5 meters
R2 IB2 to Rack 4	R2-IB2-P10A to R4-U1-P7B R2-IB2-P10B to R4-U1-P8B	5 meters
R2 IB2 to Rack 5	R2-IB2-P11A to R5-U1-P9B	10 meters

Table 6-12 shows the cable connections for the third spine switch (R3-U1) when cabling five full racks together:

Table 6-12 Leaf Switch Connections for the Third Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 within Rack 3	R3-IB3-P8A to R3-U1-P3A R3-IB3-P8B to R3-U1-P4A	3 meters
R3 IB3 to Rack 1	R3-IB3-P11A to R1-U1-P9A	5 meters
R3 IB3 to Rack 2	R3-IB3-P11B to R2-U1-P10A	5 meters
R3 IB3 to Rack 4	R3-IB3-P9A to R4-U1-P5A R3-IB3-P9B to R4-U1-P6A	5 meters
R3 IB3 to Rack 5	R3-IB3-P10A to R5-U1-P7A R3-IB3-P10B to R5-U1-P8A	5 meters
R3 IB2 within Rack 3	R3-IB2-P8A to R3-U1-P3B R3-IB2-P8B to R3-U1-P4B	3 meters
R3 IB2 to Rack 1	R3-IB2-P11A to R1-U1-P9B	5 meters
R3 IB2 to Rack 2	R3-IB2-P11B to R2-U1-P10B	5 meters
R3 IB2 to Rack 4	R3-IB2-P9A to R4-U1-P5B R3-IB2-P9B to R4-U1-P6B	5 meters
R3 IB2 to Rack 5	R3-IB2-P10A to R5-U1-P7B R3-IB2-P10B to R5-U1-P8B	5 meters

Table 6-13 shows the cable connections for the fourth spine switch (R4-U1) when cabling five full racks together:

Table 6-13 Leaf Switch Connections for the Fourth Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R4 IB3 within Rack 4	R4-IB3-P8A to R4-U1-P3A R4-IB3-P8B to R4-U1-P4A	3 meters
R4 IB3 to Rack 1	R4-IB3-P10A to R1-U1-P7A R4-IB3-P10B to R1-U1-P8A	10 meters
R4 IB3 to Rack 2	R4-IB3-P11A to R2-U1-P9A	5 meters
R4 IB3 to Rack 3	R4-IB3-P11B to R3-U1-P10A	5 meters
R4 IB3 to Rack 5	R4-IB3-P9A to R5-U1-P5A R4-IB3-P9B to R5-U1-P6A	5 meters
R4 IB2 within Rack 4	R4-IB2-P8A to R4-U1-P3B R4-IB2-P8B to R4-U1-P4B	3 meters
R4 IB2 to Rack 1	R4-IB2-P10A to R1-U1-P7B R4-IB2-P10B to R1-U1-P8B	10 meters
R4 IB2 to Rack 2	R4-IB2-P11A to R2-U1-P9B	5 meters
R4 IB2 to Rack 3	R4-IB2-P11B to R3-U1-P10B	5 meters
R4 IB2 to Rack 5	R4-IB2-P9A to R5-U1-P5B R4-IB2-P9B to R5-U1-P6B	5 meters

Table 6-14 shows the cable connections for the fifth spine switch (R5-U1) when cabling five full racks together:

Table 6-14 Leaf Switch Connections for the Fifth Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R5 IB3 within Rack 5	R5-IB3-P8A to R5-U1-P3A R5-IB3-P8B to R5-U1-P4A	3 meters
R5 IB3 to Rack 1	R5-IB3-P9A to R1-U1-P5A R5-IB3-P9B to R1-U1-P6A	10 meters
R5 IB3 to Rack 2	R5-IB3-P10A to R2-U1-P7A R5-IB3-P10B to R2-U1-P8A	10 meters
R5 IB3 to Rack 3	R5-IB3-P11A to R3-U1-P9A	5 meters
R5 IB3 to Rack 4	R5-IB3-P11B to R4-U1-P10A	5 meters
R5 IB2 within Rack 5	R5-IB2-P8A to R5-U1-P3B R5-IB2-P8B to R5-U1-P4B	3 meters
R5 IB2 to Rack 1	R5-IB2-P9A to R1-U1-P5B R5-IB2-P9B to R1-U1-P6B	10 meters
R5 IB2 to Rack 2	R5-IB2-P10A to R2-U1-P7B R5-IB2-P10B to R2-U1-P8B	10 meters
R5 IB2 to Rack 3	R5-IB2-P11A to R3-U1-P9B	5 meters

Table 6-14 (Cont.) Leaf Switch Connections for the Fifth Rack in a Five-Rack System

Leaf Switch	Connection	Cable Length
R5 IB2 to Rack 4	R5-IB2-P11B to R4-U1-P10B	5 meters

6.6 Six-Rack Cabling with InfiniBand Network Fabric

Note the following when cabling six racks together with InfiniBand Network Fabric:

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following tables, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X2-8 and later 8-socket racks, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later two-socket racks, or Oracle Exadata Database Machine X3-2 Full Racks and Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the tables assume the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Up to 100 meters is supported.
- Only optical cables are supported for lengths greater than 5 meters.

[Table 6-15](#) shows the cable connections for the first spine switch (R1-U1) when cabling six racks together:

Table 6-15 Leaf Switch Connections for the First Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A R1-IB3-P8B to R1-U1-P4A	3 meters
R1 IB3 to Rack 2	R1-IB3-P9A to R2-U1-P5A R1-IB3-P9B to R2-U1-P6A	5 meters
R1 IB3 to Rack 3	R1-IB3-P10A to R3-U1-P7A	5 meters
R1 IB3 to Rack 4	R1-IB3-P10B to R4-U1-P8A	10 meters
R1 IB3 to Rack 5	R1-IB3-P11A to R5-U1-P9A	10 meters
R1 IB3 to Rack 6	R1-IB3-P11B to R6-U1-P10A	10 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B R1-IB2-P8B to R1-U1-P4B	3 meters
R1 IB2 to Rack 2	R1-IB2-P9A to R2-U1-P5B R1-IB2-P9B to R2-U1-P6B	5 meters
R1 IB2 to Rack 3	R1-IB2-P10A to R3-U1-P7B	5 meters

Table 6-15 (Cont.) Leaf Switch Connections for the First Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R1 IB2 to Rack 4	R1-IB2-P10B to R4-U1-P8B	10 meters
R1 IB2 to Rack 5	R1-IB2-P11A to R5-U1-P9B	10 meters
R1 IB2 to Rack 6	R1-IB2-P11B to R6-U1-P10B	10 meters

[Table 6-16](#) shows the cable connections for the second spine switch (R2-U1) when cabling six full racks together:

Table 6-16 Leaf Switch Connections for the Second Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A R2-IB3-P8B to R2-U1-P4A	3 meters
R2 IB3 to Rack 1	R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB3 to Rack 3	R2-IB3-P9A to R3-U1-P5A R2-IB3-P9B to R3-U1-P6A	5 meters
R2 IB3 to Rack 4	R2-IB3-P10A to R4-U1-P7A	5 meters
R2 IB3 to Rack 5	R2-IB3-P10B to R5-U1-P8A	10 meters
R2 IB3 to Rack 6	R2-IB3-P11A to R6-U1-P9A	10 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B R2-IB2-P8B to R2-U1-P4B	3 meters
R2 IB2 to Rack 1	R2-IB2-P11B to R1-U1-P10B	5 meters
R2 IB2 to Rack 3	R2-IB2-P9A to R3-U1-P5B R2-IB2-P9B to R3-U1-P6B	5 meters
R2 IB2 to Rack 4	R2-IB2-P10A to R4-U1-P7B	5 meters
R2 IB2 to Rack 5	R2-IB2-P10B to R5-U1-P8B	10 meters
R2 IB2 to Rack 6	R2-IB2-P11A to R6-U1-P9B	10 meters

[Table 6-17](#) shows the cable connections for the third spine switch (R3-U1) when cabling six full racks together:

Table 6-17 Leaf Switch Connections for the Third Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 within Rack 3	R3-IB3-P8A to R3-U1-P3A R3-IB3-P8B to R3-U1-P4A	3 meters
R3 IB3 to Rack 1	R3-IB3-P11A to R1-U1-P9A	5 meters
R3 IB3 to Rack 2	R3-IB3-P11B to R2-U1-P10A	5 meters
R3 IB3 to Rack 4	R3-IB3-P9A to R4-U1-P5A R3-IB3-P9B to R4-U1-P6A	5 meters

Table 6-17 (Cont.) Leaf Switch Connections for the Third Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 to Rack 5	R3-IB3-P10A to R5-U1-P7A	5 meters
R3 IB3 to Rack 6	R3-IB3-P10B to R6-U1-P8A	10 meters
R3 IB2 within Rack 3	R3-IB2-P8A to R3-U1-P3B R3-IB2-P8B to R3-U1-P4B	3 meters
R3 IB2 to Rack 1	R3-IB2-P11A to R1-U1-P9B	5 meters
R3 IB2 to Rack 2	R3-IB2-P11B to R2-U1-P10B	5 meters
R3 IB2 to Rack 4	R3-IB2-P9A to R4-U1-P5B R3-IB2-P9B to R4-U1-P6B	5 meters
R3 IB2 to Rack 5	R3-IB2-P10A to R5-U1-P7B	5 meters
R3 IB2 to Rack 6	R3-IB2-P10B to R6-U1-P8B	10 meters

Table 6-18 shows the cable connections for the fourth spine switch (R4-U1) when cabling six full racks together:

Table 6-18 Leaf Switch Connections for the Fourth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R4 IB3 within Rack 4	R4-IB3-P8A to R4-U1-P3A R4-IB3-P8B to R4-U1-P4A	3 meters
R4 IB3 to Rack 1	R4-IB3-P10B to R1-U1-P8A	10 meters
R4 IB3 to Rack 2	R4-IB3-P11A to R2-U1-P9A	5 meters
R4 IB3 to Rack 3	R4-IB3-P11B to R3-U1-P10A	5 meters
R4 IB3 to Rack 5	R4-IB3-P9A to R5-U1-P5A R4-IB3-P9B to R5-U1-P6A	5 meters
R4 IB3 to Rack 6	R4-IB3-P10A to R6-U1-P7A	5 meters
R4 IB2 within Rack 4	R4-IB2-P8A to R4-U1-P3B R4-IB2-P8B to R4-U1-P4B	3 meters
R4 IB2 to Rack 1	R4-IB2-P10B to R1-U1-P8B	10 meters
R4 IB2 to Rack 2	R4-IB2-P11A to R2-U1-P9B	5 meters
R4 IB2 to Rack 3	R4-IB2-P11B to R3-U1-P10B	5 meters
R4 IB2 to Rack 5	R4-IB2-P9A to R5-U1-P5B R4-IB2-P9B to R5-U1-P6B	5 meters
R4 IB2 to Rack 6	R4-IB2-P10A to R6-U1-P7B	5 meters

Table 6-19 shows the cable connections for the fifth spine switch (R5-U1) when cabling six full racks together:

Table 6-19 Leaf Switch Connections for the Fifth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R5 IB3 within Rack 5	R5-IB3-P8A to R5-U1-P3A R5-IB3-P8B to R5-U1-P4A	3 meters
R5 IB3 to Rack 1	R5-IB3-P10A to R1-U1-P7A	10 meters
R5 IB3 to Rack 2	R5-IB3-P10B to R2-U1-P8A	10 meters
R5 IB3 to Rack 3	R5-IB3-P11A to R3-U1-P9A	5 meters
R5 IB3 to Rack 4	R5-IB3-P11B to R4-U1-P10A	5 meters
R5 IB3 to Rack 6	R5-IB3-P9A to R6-U1-P5A R5-IB3-P9B to R6-U1-P6A	5 meters
R5 IB2 within Rack 5	R5-IB2-P8A to R5-U1-P3B R5-IB2-P8B to R5-U1-P4B	3 meters
R5 IB2 to Rack 1	R5-IB2-P10A to R1-U1-P7B	10 meters
R5 IB2 to Rack 2	R5-IB2-P10B to R2-U1-P8B	10 meters
R5 IB2 to Rack 3	R5-IB2-P11A to R3-U1-P9B	5 meters
R5 IB2 to Rack 4	R5-IB2-P11B to R4-U1-P10B	5 meters
R5 IB2 to Rack 6	R5-IB2-P9A to R6-U1-P5B R5-IB2-P9B to R6-U1-P6B	5 meters

Table 6-20 shows the cable connections for the sixth spine switch (R6-U1) when cabling six full racks together:

Table 6-20 Leaf Switch Connections for the Sixth Rack in a Six-Rack System

Leaf Switch	Connection	Cable Length
R6 IB3 within Rack 6	R6-IB3-P8A to R6-U1-P3A R6-IB3-P8B to R6-U1-P4A	3 meters
R6 IB3 to Rack 1	R6-IB3-P9A to R1-U1-P5A R6-IB3-P9B to R1-U1-P6A	10 meters
R6 IB3 to Rack 2	R6-IB3-P10A to R2-U1-P7A	10 meters
R6 IB3 to Rack 3	R6-IB3-P10B to R3-U1-P8A	10 meters
R6 IB3 to Rack 4	R6-IB3-P11A to R4-U1-P9A	5 meters
R6 IB3 to Rack 5	R6-IB3-P11B to R5-U1-P10A	5 meters
R6 IB2 within Rack 6	R6-IB2-P8A to R6-U1-P3B R6-IB2-P8B to R6-U1-P4B	3 meters
R6 IB2 to Rack 1	R6-IB2-P9A to R1-U1-P5B R6-IB2-P9B to R1-U1-P6B	10 meters
R6 IB2 to Rack 2	R6-IB2-P10A to R2-U1-P7B	10 meters
R6 IB2 to Rack 3	R6-IB2-P10B to R3-U1-P8B	10 meters
R6 IB2 to Rack 4	R6-IB2-P11A to R4-U1-P9B	5 meters
R6 IB2 to Rack 5	R6-IB2-P11B to R5-U1-P10B	5 meters

6.7 Seven-Rack Cabling with InfiniBand Network Fabric

Note the following when cabling seven racks together with InfiniBand Network Fabric:

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following tables, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X2-8 and later racks, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later racks, or Oracle Exadata Database Machine X3-2 Full Racks and Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the tables assume the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Up to 100 meters is supported.
- Only optical cables are supported for lengths greater than 5 meters.

Table 6-21 shows the cable connections for the first spine switch (R1-U1) when cabling seven racks together:

Table 6-21 Leaf Switch Connections for the First Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A R1-IB3-P8B to R1-U1-P4A	3 meters
R1 IB3 to Rack 2	R1-IB3-P9A to R2-U1-P5A	5 meters
R1 IB3 to Rack 3	R1-IB3-P9B to R3-U1-P6A	5 meters
R1 IB3 to Rack 4	R1-IB3-P10A to R4-U1-P7A	10 meters
R1 IB3 to Rack 5	R1-IB3-P10B to R5-U1-P8A	10 meters
R1 IB3 to Rack 6	R1-IB3-P11A to R6-U1-P9A	10 meters
R1 IB3 to Rack 7	R1-IB3-P11B to R7-U1-P10A	10 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B R1-IB2-P8B to R1-U1-P4B	3 meters
R1 IB2 to Rack 2	R1-IB2-P9A to R2-U1-P5B	5 meters
R1 IB2 to Rack 3	R1-IB2-P9B to R3-U1-P6B	5 meters
R1 IB2 to Rack 4	R1-IB2-P10A to R4-U1-P7B	10 meters
R1 IB2 to Rack 5	R1-IB2-P10B to R5-U1-P8B	10 meters
R1 IB2 to Rack 6	R1-IB2-P11A to R6-U1-P9B	10 meters
R1 IB2 to Rack 7	R1-IB2-P11B to R7-U1-P10B	10 meters

Table 6-22 shows the cable connections for the second spine switch (R2-U1) when cabling seven full racks together:

Table 6-22 Leaf Switch Connections for the Second Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A R2-IB3-P8B to R2-U1-P4A	3 meters
R2 IB3 to Rack 1	R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB3 to Rack 3	R2-IB3-P9A to R3-U1-P5A	5 meters
R2 IB3 to Rack 4	R2-IB3-P9B to R4-U1-P6A	5 meters
R2 IB3 to Rack 5	R2-IB3-P10A to R5-U1-P7A	10 meters
R2 IB3 to Rack 6	R2-IB3-P10B to R6-U1-P8A	10 meters
R2 IB3 to Rack 7	R2-IB3-P11A to R7-U1-P9A	10 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B R2-IB2-P8B to R2-U1-P4B	3 meters
R2 IB2 to Rack 1	R2-IB2-P11B to R1-U1-P10B	5 meters
R2 IB2 to Rack 3	R2-IB2-P9A to R3-U1-P5B	5 meters
R2 IB2 to Rack 4	R2-IB2-P9B to R4-U1-P6B	5 meters
R2 IB2 to Rack 5	R2-IB2-P10A to R5-U1-P7B	10 meters
R2 IB2 to Rack 6	R2-IB2-P10B to R6-U1-P8B	10 meters
R2 IB2 to Rack 7	R2-IB2-P11A to R7-U1-P9B	10 meters

Table 6-23 shows the cable connections for the third spine switch (R3-U1) when cabling seven full racks together:

Table 6-23 Leaf Switch Connections for the Third Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 within Rack 3	R3-IB3-P8A to R3-U1-P3A R3-IB3-P8B to R3-U1-P4A	3 meters
R3 IB3 to Rack 1	R3-IB3-P11A to R1-U1-P9A	5 meters
R3 IB3 to Rack 2	R3-IB3-P11B to R2-U1-P10A	5 meters
R3 IB3 to Rack 4	R3-IB3-P9A to R4-U1-P5A	5 meters
R3 IB3 to Rack 5	R3-IB3-P9B to R5-U1-P6A	5 meters
R3 IB3 to Rack 6	R3-IB3-P10A to R6-U1-P7A	10 meters
R3 IB3 to Rack 7	R3-IB3-P10B to R7-U1-P8A	10 meters
R3 IB2 within Rack 3	R3-IB2-P8A to R3-U1-P3B R3-IB2-P8B to R3-U1-P4B	3 meters
R3 IB2 to Rack 1	R3-IB2-P11A to R1-U1-P9B	5 meters
R3 IB2 to Rack 2	R3-IB2-P11B to R2-U1-P10B	5 meters
R3 IB2 to Rack 4	R3-IB2-P9A to R4-U1-P5B	5 meters
R3 IB2 to Rack 5	R3-IB2-P9B to R5-U1-P6B	5 meters

Table 6-23 (Cont.) Leaf Switch Connections for the Third Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R3 IB2 to Rack 6	R3-IB2-P10A to R6-U1-P7B	10 meters
R3 IB2 to Rack 7	R3-IB2-P10B to R7-U1-P8B	10 meters

Table 6-24 shows the cable connections for the fourth spine switch (R4-U1) when cabling seven full racks together:

Table 6-24 Leaf Switch Connections for the Fourth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R4 IB3 within Rack 4	R4-IB3-P8A to R4-U1-P3A R4-IB3-P8B to R4-U1-P4A	3 meters
R4 IB3 to Rack 1	R4-IB3-P10B to R1-U1-P8A	10 meters
R4 IB3 to Rack 2	R4-IB3-P11A to R2-U1-P9A	5 meters
R4 IB3 to Rack 3	R4-IB3-P11B to R3-U1-P10A	5 meters
R4 IB3 to Rack 5	R4-IB3-P9A to R5-U1-P5A	5 meters
R4 IB3 to Rack 6	R4-IB3-P9B to R6-U1-P6A	5 meters
R4 IB3 to Rack 7	R4-IB3-P10A to R7-U1-P7A	10 meters
R4 IB2 within Rack 4	R4-IB2-P8A to R4-U1-P3B R4-IB2-P8B to R4-U1-P4B	3 meters
R4 IB2 to Rack 1	R4-IB2-P10B to R1-U1-P8B	10 meters
R4 IB2 to Rack 2	R4-IB2-P11A to R2-U1-P9B	5 meters
R4 IB2 to Rack 3	R4-IB2-P11B to R3-U1-P10B	5 meters
R4 IB2 to Rack 5	R4-IB2-P9A to R5-U1-P5B	5 meters
R4 IB2 to Rack 6	R4-IB2-P9B to R6-U1-P6B	5 meters
R4 IB2 to Rack 7	R4-IB2-P10A to R7-U1-P7B	10 meters

Table 6-25 shows the cable connections for the fifth spine switch (R5-U1) when cabling seven full racks together:

Table 6-25 Leaf Switch Connections for the Fifth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R5 IB3 within Rack 5	R5-IB3-P8A to R5-U1-P3A R5-IB3-P8B to R5-U1-P4A	3 meters
R5 IB3 to Rack 1	R5-IB3-P10A to R1-U1-P7A	10 meters
R5 IB3 to Rack 2	R5-IB3-P10B to R2-U1-P8A	10 meters
R5 IB3 to Rack 3	R5-IB3-P11A to R3-U1-P9A	5 meters
R5 IB3 to Rack 4	R5-IB3-P11B to R4-U1-P10A	5 meters

Table 6-25 (Cont.) Leaf Switch Connections for the Fifth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R5 IB3 to Rack 6	R5-IB3-P9A to R6-U1-P5A	5 meters
R5 IB3 to Rack 7	R5-IB3-P9B to R7-U1-P6A	5 meters
R5 IB2 within Rack 5	R5-IB2-P8A to R5-U1-P3B R5-IB2-P8B to R5-U1-P4B	3 meters
R5 IB2 to Rack 1	R5-IB2-P10A to R1-U1-P7B	10 meters
R5 IB2 to Rack 2	R5-IB2-P10B to R2-U1-P8B	10 meters
R5 IB2 to Rack 3	R5-IB2-P11A to R3-U1-P9B	5 meters
R5 IB2 to Rack 4	R5-IB2-P11B to R4-U1-P10B	5 meters
R5 IB2 to Rack 6	R5-IB2-P9A to R6-U1-P5B	5 meters
R5 IB2 to Rack 7	R5-IB2-P9B to R7-U1-P6B	5 meters

[Table 6-26](#) shows the cable connections for the sixth spine switch (R6-U1) when cabling seven full racks together:

Table 6-26 Leaf Switch Connections for the Sixth Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R6 IB3 within Rack 6	R6-IB3-P8A to R6-U1-P3A R6-IB3-P8B to R6-U1-P4A	3 meters
R6 IB3 to Rack 1	R6-IB3-P9B to R1-U1-P6A	10 meters
R6 IB3 to Rack 2	R6-IB3-P10A to R2-U1-P7A	10 meters
R6 IB3 to Rack 3	R6-IB3-P10B to R3-U1-P8A	10 meters
R6 IB3 to Rack 4	R6-IB3-P11A to R4-U1-P9A	5 meters
R6 IB3 to Rack 5	R6-IB3-P11B to R5-U1-P10A	5 meters
R6 IB3 to Rack 7	R6-IB3-P9A to R7-U1-P5A	5 meters
R6 IB2 within Rack 6	R6-IB2-P8A to R6-U1-P3B R6-IB2-P8B to R6-U1-P4B	3 meters
R6 IB2 to Rack 1	R6-IB2-P9B to R1-U1-P6B	10 meters
R6 IB2 to Rack 2	R6-IB2-P10A to R2-U1-P7B	10 meters
R6 IB2 to Rack 3	R6-IB2-P10B to R3-U1-P8B	10 meters
R6 IB2 to Rack 4	R6-IB2-P11A to R4-U1-P9B	5 meters
R6 IB2 to Rack 5	R6-IB2-P11B to R5-U1-P10B	5 meters
R6 IB2 to Rack 7	R6-IB2-P9A to R7-U1-P5B	5 meters

[Table 6-27](#) shows the cable connections for the seventh spine switch (R7-U1) when cabling seven full racks together:

Table 6-27 Leaf Switch Connections for the Seventh Rack in a Seven-Rack System

Leaf Switch	Connection	Cable Length
R7 IB3 within Rack 7	R7-IB3-P8A to R7-U1-P3A R7-IB3-P8B to R7-U1-P4A	3 meters
R7 IB3 to Rack 1	R7-IB3-P9A to R1-U1-P5A	10 meters
R7 IB3 to Rack 2	R7-IB3-P9B to R2-U1-P6A	10 meters
R7 IB3 to Rack 3	R7-IB3-P10A to R3-U1-P7A	10 meters
R7 IB3 to Rack 4	R7-IB3-P10B to R4-U1-P8A	10 meters
R7 IB3 to Rack 5	R7-IB3-P11A to R5-U1-P9A	5 meters
R7 IB3 to Rack 6	R7-IB3-P11B to R6-U1-P10A	5 meters
R7 IB2 within Rack 7	R7-IB2-P8A to R7-U1-P3B R7-IB2-P8B to R7-U1-P4B	3 meters
R7 IB2 to Rack 1	R7-IB2-P9A to R1-U1-P5B	10 meters
R7 IB2 to Rack 2	R7-IB2-P9B to R2-U1-P6B	10 meters
R7 IB2 to Rack 3	R7-IB2-P10A to R3-U1-P7B	10 meters
R7 IB2 to Rack 4	R7-IB2-P10B to R4-U1-P8B	10 meters
R7 IB2 to Rack 5	R7-IB2-P11A to R5-U1-P9B	5 meters
R7 IB2 to Rack 6	R7-IB2-P11B to R6-U1-P10B	5 meters

6.8 Eight-Rack Cabling with InfiniBand Network Fabric

Note the following when cabling eight racks together with InfiniBand Network Fabric:

- The spine switch is in U1 (also referred to as IB1) for all racks.
- In the following tables, the leaf switches are referred to as IB2 and IB3. Use the rack unit that is appropriate for your system.
- In Oracle Exadata Database Machine X2-2 racks and Oracle Exadata Storage Expansion Racks with Exadata Storage Server with Sun Fire X4270 M2 Servers, the leaf switches are in U20 and U24, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X2-8 and later racks, the leaf switches are in U21 and U23, also referred to as IB2 and IB3, respectively.
- In Oracle Exadata Database Machine X4-2 and later racks, or Oracle Exadata Database Machine X3-2 Full Racks and Oracle Exadata Storage Expansion Rack X3-2 racks, the leaf switches are in U20 and U22, also referred to as IB2 and IB3, respectively.
- The cable lengths shown in the tables assume the racks are adjacent to each other. If the racks are not adjacent or use overhead cabling trays, then they may require longer cables lengths. Up to 100 meters is supported.
- Only optical cables are supported for lengths greater than 5 meters.

Table 6-28 shows the cable connections for the first spine switch (R1-U1) when cabling eight racks together:

Table 6-28 Leaf Switch Connections for the First Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R1 IB3 within Rack 1	R1-IB3-P8A to R1-U1-P3A	3 meters
R1 IB3 to Rack 2	R1-IB3-P8B to R2-U1-P4A	5 meters
R1 IB3 to Rack 3	R1-IB3-P9A to R3-U1-P5A	5 meters
R1 IB3 to Rack 4	R1-IB3-P9B to R4-U1-P6A	10 meters
R1 IB3 to Rack 5	R1-IB3-P10A to R5-U1-P7A	10 meters
R1 IB3 to Rack 6	R1-IB3-P10B to R6-U1-P8A	10 meters
R1 IB3 to Rack 7	R1-IB3-P11A to R7-U1-P9A	10 meters
R1 IB3 to Rack 8	R1-IB3-P11B to R8-U1-P10A	10 meters
R1 IB2 within Rack 1	R1-IB2-P8A to R1-U1-P3B	3 meters
R1 IB2 to Rack 2	R1-IB2-P8B to R2-U1-P4B	5 meters
R1 IB2 to Rack 3	R1-IB2-P9A to R3-U1-P5B	5 meters
R1 IB2 to Rack 4	R1-IB2-P9B to R4-U1-P6B	10 meters
R1 IB2 to Rack 5	R1-IB2-P10A to R5-U1-P7B	10 meters
R1 IB2 to Rack 6	R1-IB2-P10B to R6-U1-P8B	10 meters
R1 IB2 to Rack 7	R1-IB2-P11A to R7-U1-P9B	10 meters
R1 IB2 to Rack 8	R1-IB2-P11B to R8-U1-P10B	10 meters

Table 6-29 shows the cable connections for the second spine switch (R2-U1) when cabling eight full racks together:

Table 6-29 Leaf Switch Connections for the Second Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R2 IB3 within Rack 2	R2-IB3-P8A to R2-U1-P3A	3 meters
R2 IB3 to Rack 1	R2-IB3-P11B to R1-U1-P10A	5 meters
R2 IB3 to Rack 3	R2-IB3-P8B to R3-U1-P4A	5 meters
R2 IB3 to Rack 4	R2-IB3-P9A to R4-U1-P5A	5 meters
R2 IB3 to Rack 5	R2-IB3-P9B to R5-U1-P6A	10 meters
R2 IB3 to Rack 6	R2-IB3-P10A to R6-U1-P7A	10 meters
R2 IB3 to Rack 7	R2-IB3-P10B to R7-U1-P8A	10 meters
R2 IB3 to Rack 8	R2-IB3-P11A to R8-U1-P9A	10 meters
R2 IB2 within Rack 2	R2-IB2-P8A to R2-U1-P3B	3 meters
R2 IB2 to Rack 1	R2-IB2-P11B to R1-U1-P10B	5 meters
R2 IB2 to Rack 3	R2-IB2-P8B to R3-U1-P4B	5 meters
R2 IB2 to Rack 4	R2-IB2-P9A to R4-U1-P5B	5 meters
R2 IB2 to Rack 5	R2-IB2-P9B to R5-U1-P6B	10 meters
R2 IB2 to Rack 6	R2-IB2-P10A to R6-U1-P7B	10 meters
R2 IB2 to Rack 7	R2-IB2-P10B to R7-U1-P8B	10 meters

Table 6-29 (Cont.) Leaf Switch Connections for the Second Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R2 IB2 to Rack 8	R2-IB2-P11A to R8-U1-P9B	10 meters

Table 6-30 shows the cable connections for the third spine switch (R3-U1) when cabling eight full racks together:

Table 6-30 Leaf Switch Connections for the Third Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R3 IB3 within Rack 3	R3-IB3-P8A to R3-U1-P3A	3 meters
R3 IB3 to Rack 1	R3-IB3-P11A to R1-U1-P9A	5 meters
R3 IB3 to Rack 2	R3-IB3-P11B to R2-U1-P10A	5 meters
R3 IB3 to Rack 4	R3-IB3-P8B to R4-U1-P4A	5 meters
R3 IB3 to Rack 5	R3-IB3-P9A to R5-U1-P5A	5 meters
R3 IB3 to Rack 6	R3-IB3-P9B to R6-U1-P6A	10 meters
R3 IB3 to Rack 7	R3-IB3-P10A to R7-U1-P7A	10 meters
R3 IB3 to Rack 8	R3-IB3-P10B to R8-U1-P8A	10 meters
R3 IB2 within Rack 3	R3-IB2-P8A to R3-U1-P3B	3 meters
R3 IB2 to Rack 1	R3-IB2-P11A to R1-U1-P9B	5 meters
R3 IB2 to Rack 2	R3-IB2-P11B to R2-U1-P10B	5 meters
R3 IB2 to Rack 4	R3-IB2-P8B to R4-U1-P4B	5 meters
R3 IB2 to Rack 5	R3-IB2-P9A to R5-U1-P5B	5 meters
R3 IB2 to Rack 6	R3-IB2-P9B to R6-U1-P6B	10 meters
R3 IB2 to Rack 7	R3-IB2-P10A to R7-U1-P7B	10 meters
R3 IB2 to Rack 8	R3-IB2-P10B to R8-U1-P8B	10 meters

Table 6-31 shows the cable connections for the fourth spine switch (R4-U1) when cabling eight full racks together:

Table 6-31 Leaf Switch Connections for the Fourth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R4 IB3 within Rack 4	R4-IB3-P8A to R4-U1-P3A	3 meters
R4 IB3 to Rack 1	R4-IB3-P10B to R1-U1-P8A	10 meters
R4 IB3 to Rack 2	R4-IB3-P11A to R2-U1-P9A	5 meters
R4 IB3 to Rack 3	R4-IB3-P11B to R3-U1-P10A	5 meters
R4 IB3 to Rack 5	R4-IB3-P8B to R5-U1-P4A	5 meters
R4 IB3 to Rack 6	R4-IB3-P9A to R6-U1-P5A	5 meters
R4 IB3 to Rack 7	R4-IB3-P9B to R7-U1-P6A	10 meters

Table 6-31 (Cont.) Leaf Switch Connections for the Fourth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R4 IB3 to Rack 8	R4-IB3-P10A to R8-U1-P7A	10 meters
R4 IB2 within Rack 4	R4-IB2-P8A to R4-U1-P3B	3 meters
R4 IB2 to Rack 1	R4-IB2-P10B to R1-U1-P8B	10 meters
R4 IB2 to Rack 2	R4-IB2-P11A to R2-U1-P9B	5 meters
R4 IB2 to Rack 3	R4-IB2-P11B to R3-U1-P10B	5 meters
R4 IB2 to Rack 5	R4-IB2-P8B to R5-U1-P4B	5 meters
R4 IB2 to Rack 6	R4-IB2-P9A to R6-U1-P5B	5 meters
R4 IB2 to Rack 7	R4-IB2-P9B to R7-U1-P6B	10 meters
R4 IB2 to Rack 8	R4-IB2-P10A to R8-U1-P7B	10 meters

Table 6-32 shows the cable connections for the fifth spine switch (R5-U1) when cabling eight full racks together:

Table 6-32 Leaf Switch Connections for the Fifth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R5 IB3 within Rack 5	R5-IB3-P8A to R5-U1-P3A	3 meters
R5 IB3 to Rack 1	R5-IB3-P10A to R1-U1-P7A	10 meters
R5 IB3 to Rack 2	R5-IB3-P10B to R2-U1-P8A	10 meters
R5 IB3 to Rack 3	R5-IB3-P11A to R3-U1-P9A	5 meters
R5 IB3 to Rack 4	R5-IB3-P11B to R4-U1-P10A	5 meters
R5 IB3 to Rack 6	R5-IB3-P8B to R6-U1-P4A	5 meters
R5 IB3 to Rack 7	R5-IB3-P9A to R7-U1-P5A	5 meters
R5 IB3 to Rack 8	R5-IB3-P9B to R8-U1-P6A	10 meters
R5 IB2 within Rack 5	R5-IB2-P8A to R5-U1-P3B	3 meters
R5 IB2 to Rack 1	R5-IB2-P10A to R1-U1-P7B	10 meters
R5 IB2 to Rack 2	R5-IB2-P10B to R2-U1-P8B	10 meters
R5 IB2 to Rack 3	R5-IB2-P11A to R3-U1-P9B	5 meters
R5 IB2 to Rack 4	R5-IB2-P11B to R4-U1-P10B	5 meters
R5 IB2 to Rack 6	R5-IB2-P8B to R6-U1-P4B	5 meters
R5 IB2 to Rack 7	R5-IB2-P9A to R7-U1-P5B	5 meters
R5 IB2 to Rack 8	R5-IB2-P9B to R8-U1-P6B	10 meters

Table 6-33 shows the cable connections for the sixth spine switch (R6-U1) when cabling eight full racks together:

Table 6-33 Leaf Switch Connections for the Sixth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R6 IB3 within Rack 6	R6-IB3-P8A to R6-U1-P3A	3 meters
R6 IB3 to Rack 1	R6-IB3-P9B to R1-U1-P6A	10 meters
R6 IB3 to Rack 2	R6-IB3-P10A to R2-U1-P7A	10 meters
R6 IB3 to Rack 3	R6-IB3-P10B to R3-U1-P8A	10 meters
R6 IB3 to Rack 4	R6-IB3-P11A to R4-U1-P9A	5 meters
R6 IB3 to Rack 5	R6-IB3-P11B to R5-U1-P10A	5 meters
R6 IB3 to Rack 7	R6-IB3-P8B to R7-U1-P4A	5 meters
R6 IB3 to Rack 8	R6-IB3-P9A to R8-U1-P5A	5 meters
R6 IB2 within Rack 6	R6-IB2-P8A to R6-U1-P3B	3 meters
R6 IB2 to Rack 1	R6-IB2-P9B to R1-U1-P6B	10 meters
R6 IB2 to Rack 2	R6-IB3-P10A to R2-U1-P7B	10 meters
R6 IB2 to Rack 3	R6-IB2-P10B to R3-U1-P8B	10 meters
R6 IB2 to Rack 4	R6-IB2-P11A to R4-U1-P9B	5 meters
R6 IB2 to Rack 5	R6-IB2-P11B to R5-U1-P10B	5 meters
R6 IB2 to Rack 7	R6-IB2-P8B to R7-U1-P4B	5 meters
R6 IB2 to Rack 8	R6-IB2-P9A to R8-U1-P5B	5 meters

Table 6-34 shows the cable connections for the seventh spine switch (R7-U1) when cabling eight full racks together:

Table 6-34 Leaf Switch Connections for the Seventh Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R7 IB3 within Rack 7	R7-IB3-P8A to R7-U1-P3A	3 meters
R7 IB3 to Rack 1	R7-IB3-P9A to R1-U1-P5A	10 meters
R7 IB3 to Rack 2	R7-IB3-P9B to R2-U1-P6A	10 meters
R7 IB3 to Rack 3	R7-IB3-P10A to R3-U1-P7A	10 meters
R7 IB3 to Rack 4	R7-IB3-P10B to R4-U1-P8A	10 meters
R7 IB3 to Rack 5	R7-IB3-P11A to R5-U1-P9A	5 meters
R7 IB3 to Rack 6	R7-IB3-P11B to R6-U1-P10A	5 meters
R7 IB3 to Rack 8	R7-IB3-P8B to R8-U1-P4A	5 meters
R7 IB2 within Rack 7	R7-IB2-P8A to R7-U1-P3B	3 meters
R7 IB2 to Rack 1	R7-IB2-P9A to R1-U1-P5B	10 meters
R7 IB2 to Rack 2	R7-IB2-P9B to R2-U1-P6B	10 meters
R7 IB2 to Rack 3	R7-IB2-P10A to R3-U1-P7B	10 meters
R7 IB2 to Rack 4	R7-IB2-P10B to R4-U1-P8B	10 meters
R7 IB2 to Rack 5	R7-IB2-P11A to R5-U1-P9B	5 meters

Table 6-34 (Cont.) Leaf Switch Connections for the Seventh Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R7 IB2 to Rack 6	R7-IB2-P11B to R6-U1-P10B	5 meters
R7 IB2 to Rack 8	R7-IB2-P8B to R8-U1-P4B	5 meters

Table 6-35 shows the cable connections for the eighth spine switch (R8-U1) when cabling eight full racks together:

Table 6-35 Leaf Switch Connections for the Eighth Rack in a Eight-Rack System

Leaf Switch	Connection	Cable Length
R8 IB3 within Rack 8	R8-IB3-P8A to R8-U1-P3A	3 meters
R8 IB3 to Rack 1	R8-IB3-P8B to R1-U1-P4A	10 meters
R8 IB3 to Rack 2	R8-IB3-P9A to R2-U1-P5A	10 meters
R8 IB3 to Rack 3	R8-IB3-P9B to R3-U1-P6A	10 meters
R8 IB3 to Rack 4	R8-IB3-P10A to R4-U1-P7A	10 meters
R8 IB3 to Rack 5	R8-IB3-P10B to R5-U1-P8A	10 meters
R8 IB3 to Rack 6	R8-IB3-P11A to R6-U1-P9A	5 meters
R8 IB3 to Rack 7	R8-IB3-P11B to R7-U1-P10A	5 meters
R8 IB2 within Rack 8	R8-IB2-P8A to R8-U1-P3B	3 meters
R8 IB2 to Rack 1	R8-IB2-P8B to R1-U1-P4B	10 meters
R8 IB2 to Rack 2	R8-IB2-P9A to R2-U1-P5B	10 meters
R8 IB2 to Rack 3	R8-IB2-P9B to R3-U1-P6B	10 meters
R8 IB2 to Rack 4	R8-IB2-P10A to R4-U1-P7B	10 meters
R8 IB2 to Rack 5	R8-IB2-P10B to R5-U1-P8B	10 meters
R8 IB2 to Rack 6	R8-IB2-P11A to R6-U1-P9B	5 meters
R8 IB2 to Rack 7	R8-IB2-P11B to R7-U1-P10B	5 meters