SGI® DataRaptor™ Appliance with MarkLogic® Database

Quick Start Guide

007-5907-001

COPYRIGHT
© 2013 Silicon Graphics International Corp. All rights reserved; provided portions may be copyright in third parties, as indicated elsewhere herein. No permission is granted to copy, distribute, or create derivative works from the contents of this electronic documentation in any manner, in whole or in part, without the prior written permission of SGI.

LIMITED RIGHTS LEGEND
The software described in this document is “commercial computer software” provided with restricted rights (except as to included open/free source) as specified in the FAR 52.227-19 and/or the DFAR 227.7202, or successive sections. Use beyond license provisions is a violation of worldwide intellectual property laws, treaties and conventions. This document is provided with limited rights as defined in 52.227-14.

The electronic (software) version of this document was developed at private expense; if acquired under an agreement with the USA government or any contractor thereto, it is acquired as “commercial computer software” subject to the provisions of its applicable license agreement, as specified in (a) 48 CFR 12.212 of the FAR; or, if acquired for Department of Defense units, (b) 48 CFR 227-7202 of the DoD FAR Supplement; or sections succeeding thereto. Contractor/manufacturer is SGI, 46600 Landing Parkway, Fremont, CA 94538.

TRADEMARKS AND ATTRIBUTIONS
Silicon Graphics, SGI, the SGI logo, DataRaptor, Rackable, and Supportfolio are trademarks or registered trademarks of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries worldwide.

Intel and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

MarkLogic is a registered trademark of MarkLogic Corporation.

Red Hat and all Red Hat-based trademarks are trademarks or registered trademarks of Red Hat, Inc. in the United States and other countries.

All other trademarks mentioned herein are the property of their respective owners.

Record of Revision

Version	Description
001	January 2013 Initial printing.

Contents

About This Guide

This guide provides an overview of the SGI® DataRaptor™ with MarkLogic® Database appliance along with getting-started instructions. This guide consists of the following chapters:

- Chapter 1, “Overview,” provides an overview of the SGI DataRaptor appliance.
- Chapter 2, “Cluster Startup,” describes configuration requirements for cluster management and monitoring, getting-started instructions, and licensing.
- Chapter 3, “MarkLogic Usage Notes,” provides licensing and validation database information as well as a pointer to SGI Knowledgebase.

Audience

This guide is written for the system administrators of the DataRaptor appliance and developers. The guide assumes the reader is familiar with clusters and Big Data technology.

Related Publications

The following SGI documents are relevant to your SGI DataRaptor appliance:

- *SGI Management Center Quick Start Guide* (007-5672-xxx)
- *SGI Management Center (SMC) Installation and Configuration* (007-5643-xxx)
- *SGI Management Center (SMC) System Administrator's Guide* (007-5642-xxx)
- *SGI Rackable RP2 Standard-Depth Servers User Guide* (007-5837-xxx)
- *SGI Rackable C1110-RP6 System User Guide* (007-5843-xxx)

You can obtain SGI documentation in the following ways:

- Refer to the SGI Technical Publications Library (TPL) at http://docs.sgi.com. Various formats are available. The TPL contains the most recent and most comprehensive set of online books, man pages, and other information.

  To get the latest revision of a document on the TPL, use the core publication number as your search string. For example, use `007-1234` as your search string to get the latest version of the document with part number 007-1234-xxx.
- Refer to the SGI Supportfolio™ webpage for documents whose access require a support contract. See "Product Support" on page ix.
- You can also view man pages by typing **`man`** *<title>* on a command line.

Note: For information about third-party system components, see the documentation provided by the manufacturer/supplier.

Product Support

SGI provides a comprehensive product support and maintenance program for its products. SGI also offers services to implement and integrate Linux applications in your environment.

- Refer to http://www.sgi.com/support/
- If you are in North America, contact the Technical Assistance Center at +1 800 800 4SGI or contact your authorized service provider.
- If you are outside North America, contact the SGI subsidiary or authorized distributor in your country.

Be sure to have the following information before you call Technical Support:

- Product serial number
- Product model name and number
- Applicable error messages
- Add-on boards or hardware
- Third-party hardware or software
- Operating system type and revision level

Reader Comments

If you have comments about the technical accuracy, content, or organization of this document, contact SGI. Be sure to include the title and document number of the manual with your comments. (Online, the document number is located in the front matter of the manual. In printed manuals, the document number is located at the bottom of each page.)

You can contact SGI in any of the following ways:

- Send e-mail to the following address: techpubs@sgi.com
- Contact your customer service representative and ask that an incident be filed in the SGI incident tracking system.
- Send mail to the following address:

  SGI
  Technical Publications
  46600 Landing Parkway
  Fremont, CA 94538

SGI values your comments and will respond to them promptly.

Overview

As illustrated in Figure 1-1, the SGI DataRaptor with MarkLogic Database appliance provides a content management platform for Information Intelligence.

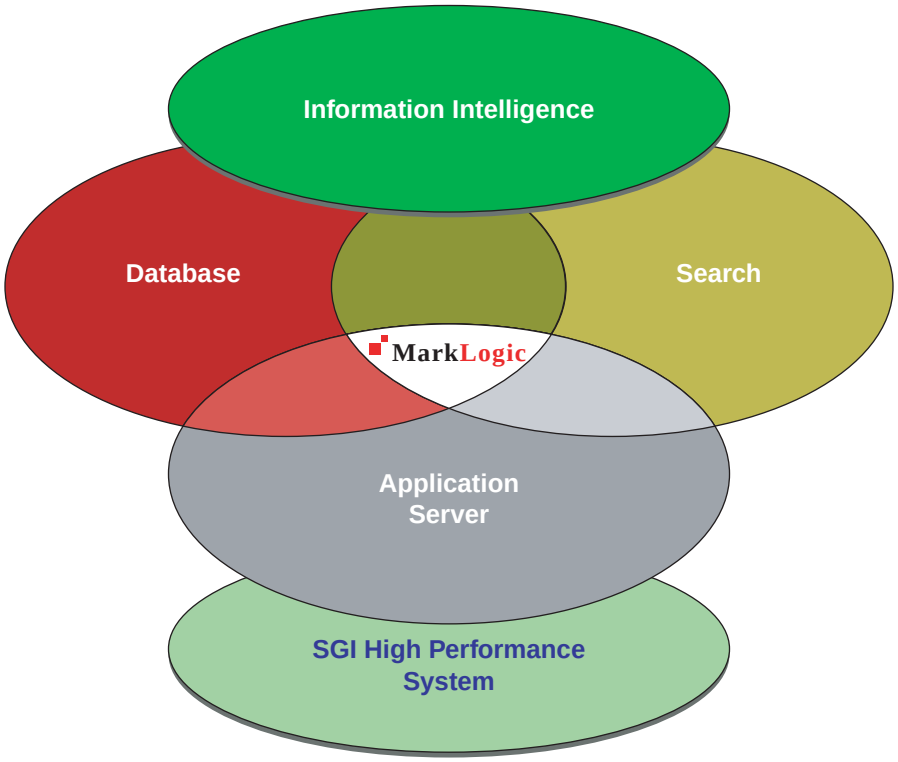

Figure 1-1 SGI DataRaptor with MarkLogic Database—Solution Stack

The appliance has the following features:

- Two appliance models:
  - SGI DataRaptor High-Performance Appliance
  - SGI DataRaptor High-Capacity Appliance
- Includes a MarkLogic NoSQL validation database.
- Delivered factory-integrated and ready-to-run
- Choice of four densely-packed rack configurations for each model:
  - One-Fourth Rack
  - Half-Rack
  - Full-Rack
  - Multi-Rack
- Ability for application developers (end users or ISVs) to test drive and then reconfigure the database per their application needs
- Can accommodate best practices and tunables for any database reconfigurations needed for user applications.

This overview describes the following components:

- “Hardware” on page 3
- “Rack Configurations” on page 6
- “Network Topology” on page 12
- “Software” on page 16

Hardware

This section describes the hardware used in the two SGI DataRaptor appliance models: first, the servers and then the network hardware.

Servers

This section describes the SGI servers that are used in the SGI DataRaptor appliance, their function in the appliance, and their specifications.

The SGI High-Performance Appliance model employs the SGI Rackable™ ISS3124-RP2 server shown in Figure 1-2.

Figure 1-2 ISS3124-RP2 Server—High-Performance Appliance Model

The SGI High-Capacity Appliance model employs the SGI Rackable ISS3112-RP2 server shown in Figure 1-3.

Figure 1-3 ISS3112-RP2 Server—High-Capacity Appliance Model

Table 1-1 describes the specifications of the servers in the high-performance model based on the Intel® Xeon® Processor E5-2600 Series.

Table 1-1 Server Specifications for the High-Performance Appliance

SGI Server	Node Function	Specifications
ISS3124-RP2	Database Node	- 2U full-depth chassis - 20x2.5” 900GB 10K rpm SAS drives in RAID 10 configuration - 4x100GB enterprise-class flash SSDs in RAID 10 configuration - Separate 80GB SSD for OS drive - Intel Xeon Processors E5-2680 (2.7GHz eight-core) - 16x 8GB 1.5v 1600 MHz DIMMs (128GB Memory) - Dual-port 10GigE
C1110-RP6	SGI Management Center Node (Admin Node)	- 1U full-depth chassis with 4x3.5” 1TB 7.2K rpm SATA drives in RAID 10 configuration - Intel Xeon Processors E5-2650 (2.0GHz eight-core) - 8x 4GB 1.5v 1600 MHz DIMMs (32GB Memory) - 3x GigE ports - Redundant power supply

Table 1-2 describes the specifications of the servers in the high-capacity model based on the Intel Xeon Processor E5-2600 Series.

Table 1-2 Server Specifications for the High-Capacity Appliance

SGI Server	Node Function	Specifications
ISS3112-RP2	Database Node	- 2U full-depth chassis - 10x3.5” 3TB 7.2K rpm SAS drives RAID 6 configuration - 2x200GB enterprise-class flash SSDs in RAID 1 configuration - Separate 80GB SSD for OS drive - Intel Xeon Processors E5-2680 (2.7GHz eight-core) - 16x 8GB 1.5v 1600 MHz DIMMs (128GB Memory) - Dual-port 10GigE
C1110-RP6	SGI Management Center Node (Admin Node)	- 1U full-depth chassis with 4x3.5” 1TB 7.2K rpm SATA drives in RAID 10 configuration - Intel Xeon Processors E5-2650 (2.0GHz eight-core) - 8x 4GB 1.5v 1600 MHz DIMMs (32GB Memory) - 3x GigE ports - Redundant power supply

Network Hardware

The following network hardware components are used in rack configurations:

- 2x 48-port 10-GigE data network switches
- 1x 48-port GigE switch for SGI Management Center network
- 2x 48-port 10-GigE spine switches with the first rack of multi-rack configurations only

Rack Configurations

The high-performance and high-capacity appliance models are available in single-rack and multi-rack configurations. This section describes the quarter-rack, half-rack, full-rack, and multi-rack configurations for both models.

Figure 1-4 shows the range of data capacity for high-performance configurations.

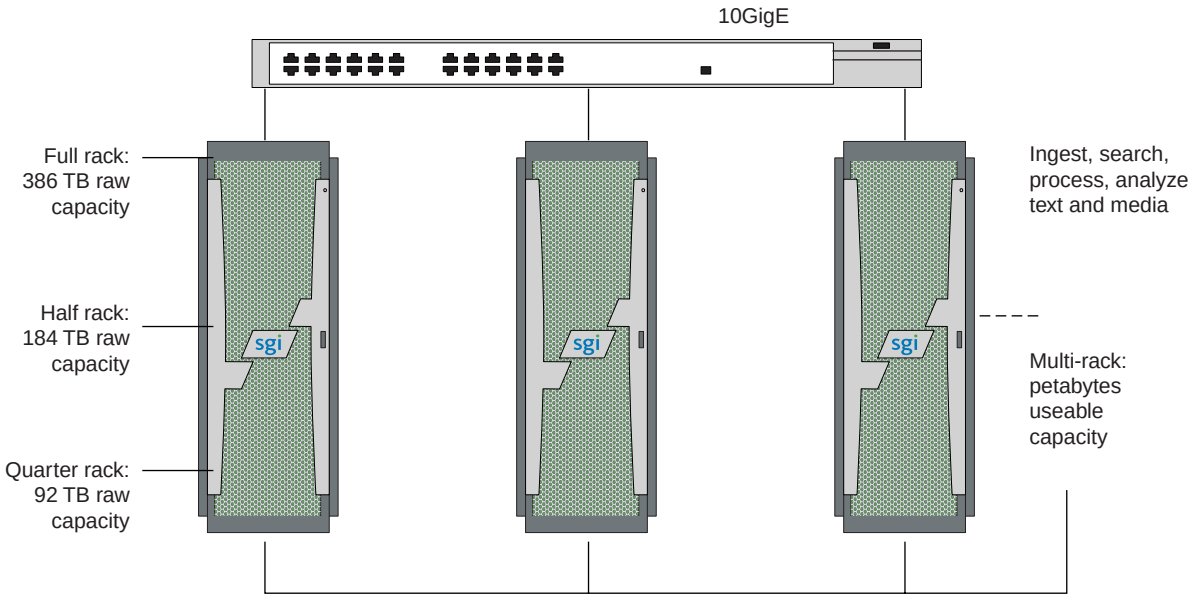

Figure 1-4 Data Capacity for Various High-Performance Rack Configurations

Figure 1-5 shows the range of data capacity for high-capacity configurations.

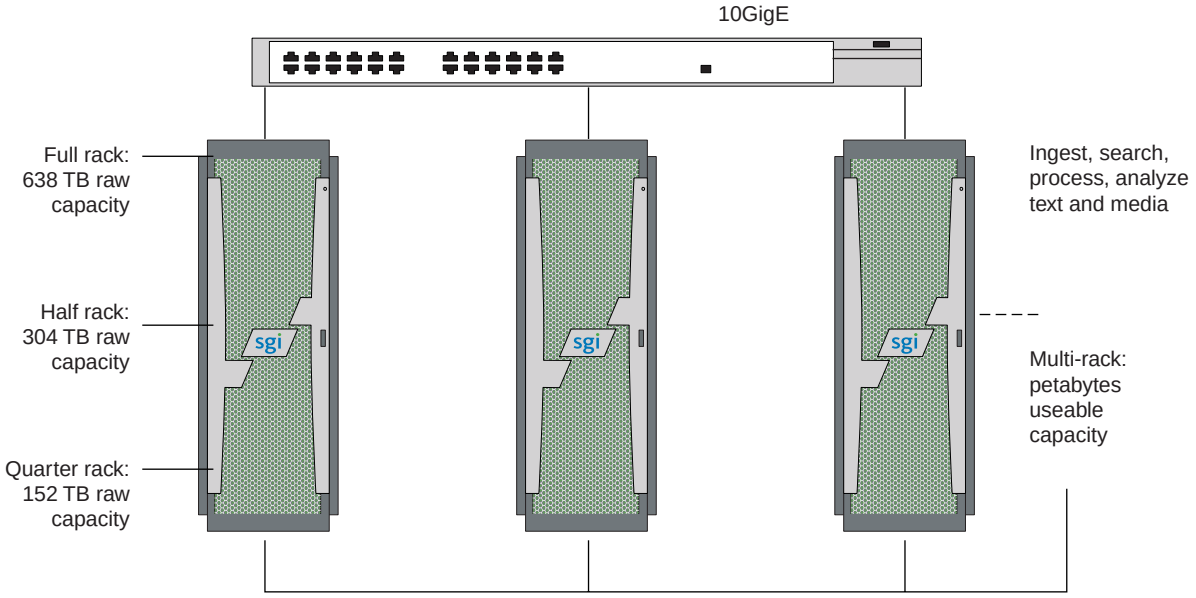

Figure 1-5 Data Capacity for Various High-Capacity Rack Configurations

Note: The following configurations shown in this section are for the high-performance appliance models. In each case, there is a similar configuration for the high-capacity appliance model.

Quarter-Rack

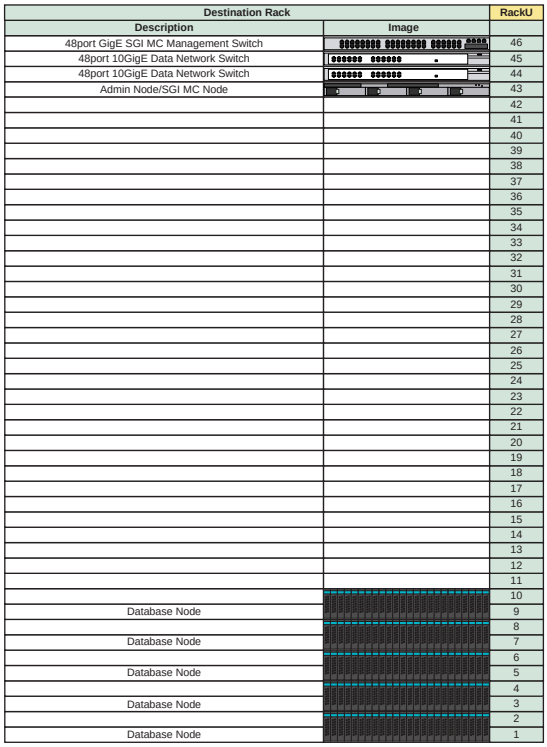

Figure 1-6 Quarter-Rack Configuration—High-Performance Appliance

Figure 1-6 describes the configuration of a quarter-rack configuration for a high-performance appliance. The rack consists of the following:

- One SGI Management Center node
- Five database nodes
- Two 48-port, 10-GigE data network switches
- One SGI Management Center network switch

Half-Rack

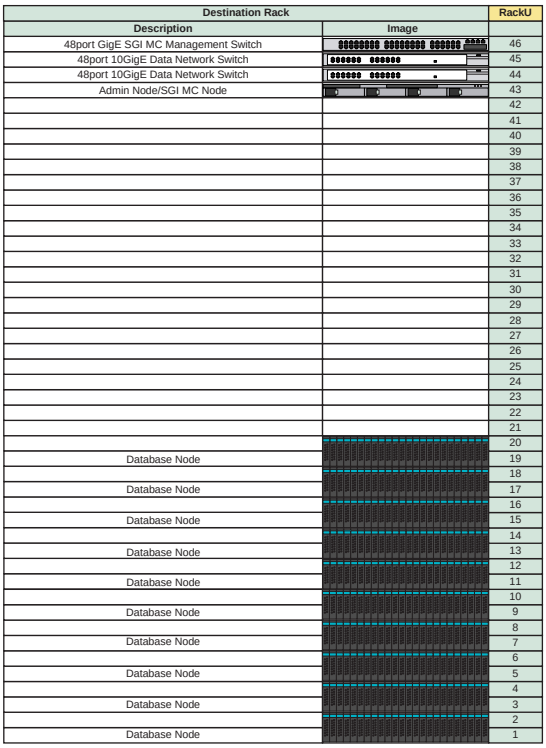

Figure 1-7 Half-Rack Configuration—High-Performance Appliance

Figure 1-7 describes the configuration of a half-rack configuration for a high-performance appliance. The rack consists of the following:

- One SGI Management Center node
- 10 Database nodes
- Two 48-port, 10-GigE data network switches
- One SGI Management Center network switch
- Two 48-port, 10-GigE data network switches

Full-Rack (46U)

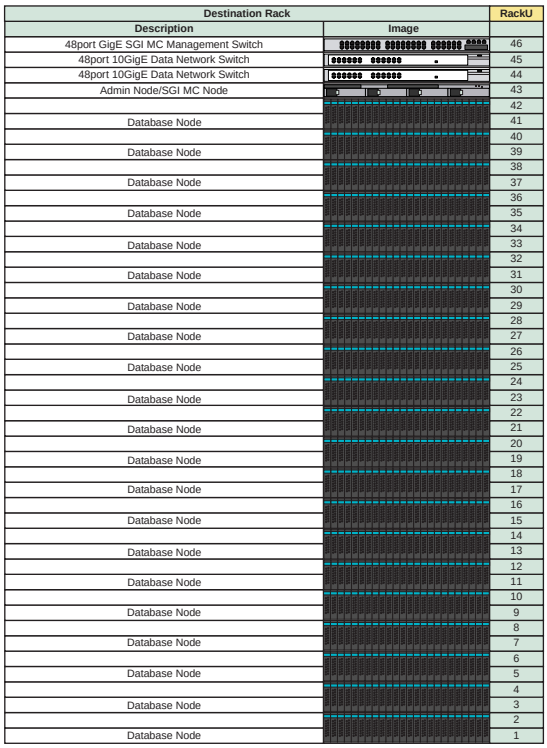

Figure 1-8 Full-Rack Configuration—High-Performance Appliance

Figure 1-8 describes the configuration of a full-rack configuration for a high-performance appliance. The rack consists of the following:

- One SGI Management Center node
- 21 Database nodes
- Two 48-port, 10-GigE data network switches
- One SGI Management Center network switch

Multi-Rack

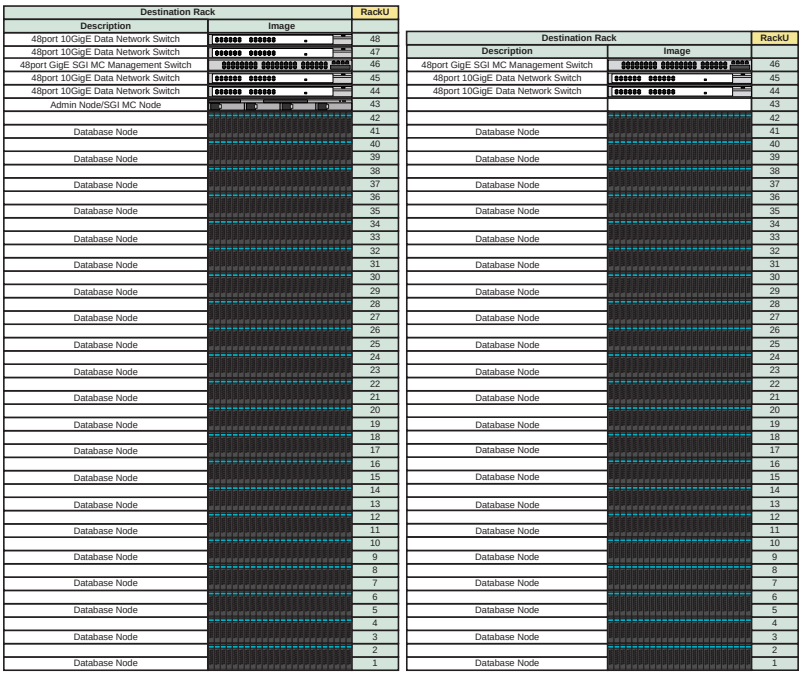

Figure 1-9 Multi-Rack (First Rack, Second Rack, and beyond)—High-Performance Appliance

Figure 1-9 describes the configuration of a multi-rack configuration for a high-performance appliance. The first rack of a multi-rack configuration consists of the following:

- One SGI Management Center node
- 21 Database nodes
- Two 48-port, 10-GigE data network switches
- One SGI Management Center network switch
- Two 48-port, 10-GigE data spine switches

The second rack (and subsequent racks) of a multi-rack configuration consists of the following:

- 21 Database nodes
- Two 48-port, 10-GigE data network switches
- One SGI Management Center network switch

Network Topology

The section illustrates the network topology from the most granular level (node level) to the top level (inter-rack level):

- "Node Level" on page 13
- "Rack Level for Single-Rack Configuration" on page 15
- "Inter-Rack Level" on page 16

Note: The network topology described in this section applies to both the high-performance and high-capacity models of the appliance.

Node Level

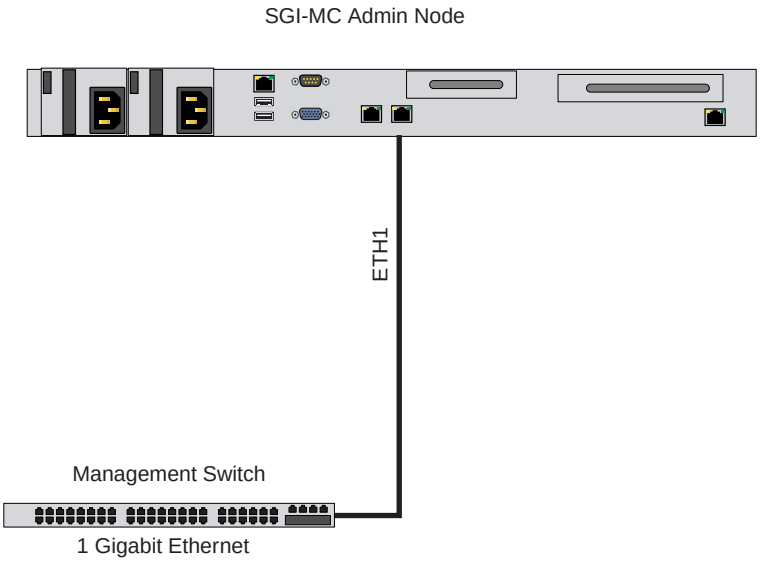

Figure 1-10 Network Topology (Node Level)—SGI Administration Node

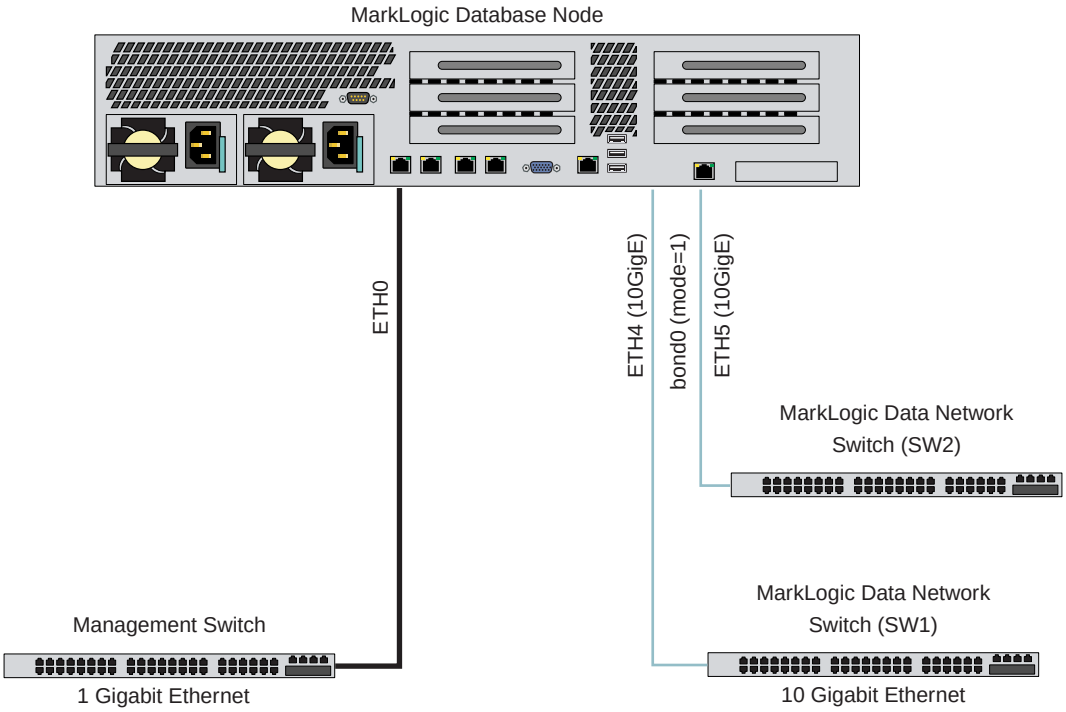

Figure 1-11 Network Topology (Node Level)—Database Server

Rack Level for Single-Rack Configuration

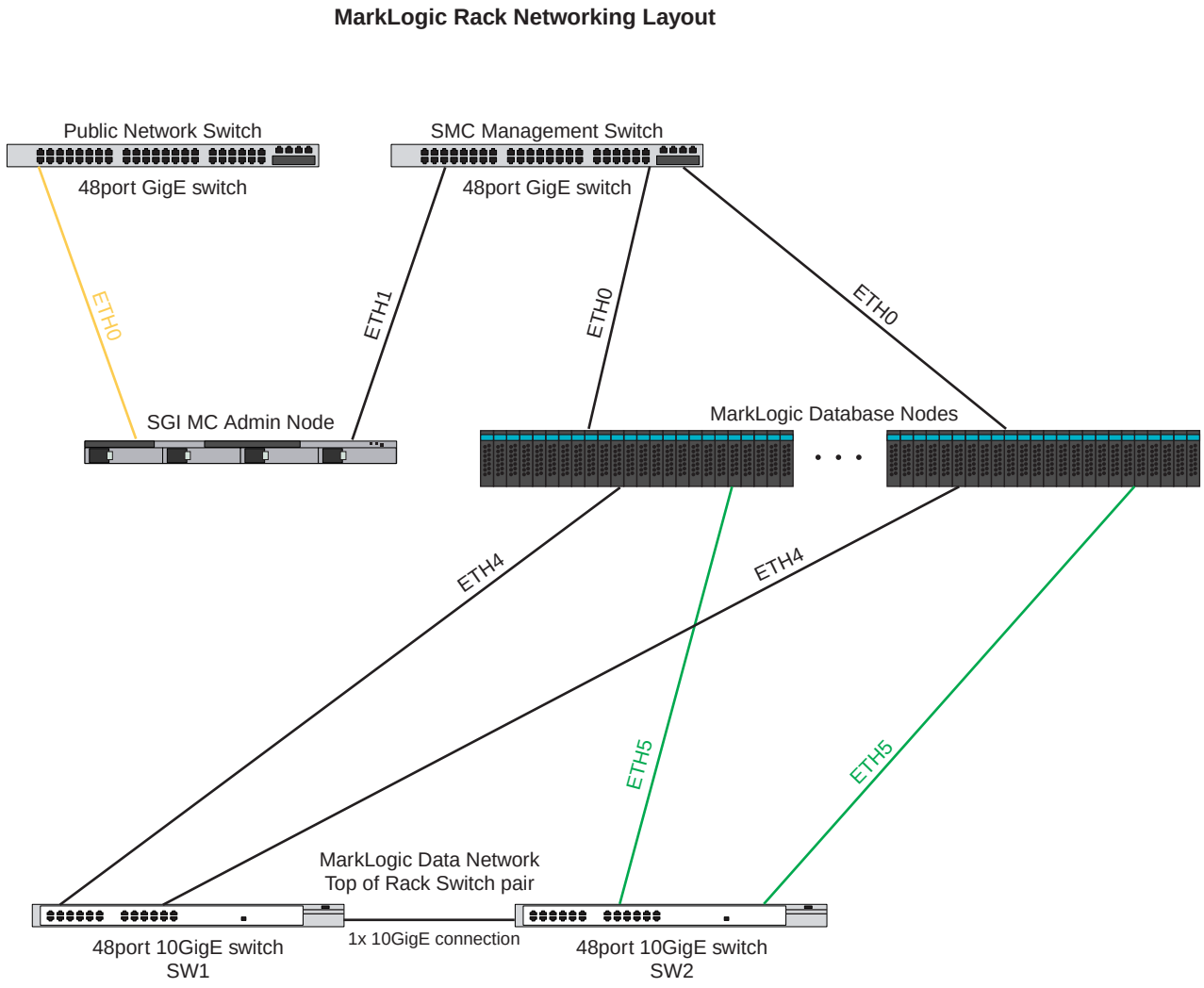

Figure 1-12 Network Topology—Rack Level for Single Rack

Inter-Rack Level

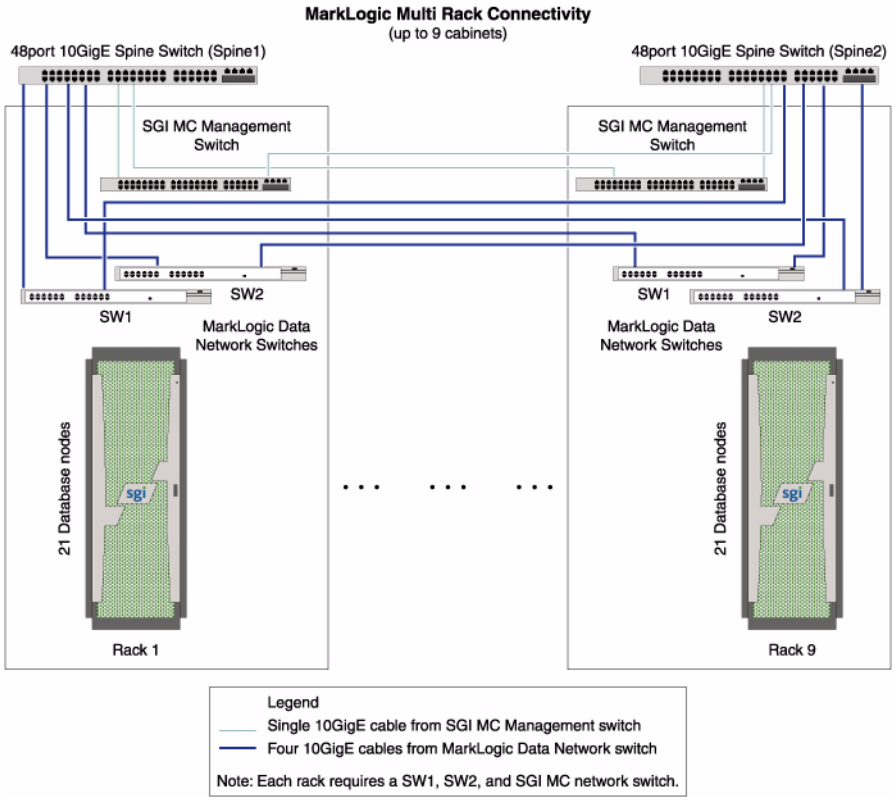

Figure 1-13 Network Topology—Inter-Rack Level

Software

The following components comprise the software stack for the SGI DataRaptor with MarkLogic Database appliance:

- Red Hat® Enterprise Linux (RHEL) 6.2 (2.6.32-220.el6.x86_64)
- MarkLogic 6 Database v6.0-1.1
- SGI Management Center 1.6

Cluster Startup

This chapter describes the broad steps for starting the SGI DataRaptor appliance:

Configuring and Starting SGI Management Center

To configure and start the SGI Management Center to monitor the DataRaptor appliance, you will need to follow the instructions in the *SGI Management Center Quick Start Guide* and appropriately configure the appliance servers described in Table 2-1.

Table 2-1 Hostnames for SGI Appliance Servers

Daemon	Management Network Hostname	Data Network Hostname
SGI Management Center	`admin`	
MarkLogic	`r`[*rack#*]`n`[*node#*]	`r`[*rack#*]`n`[*node#*]`-bond0`

Starting the Cluster for the First Time

Use the following steps to start the cluster for the first time.

1. Power on the head node of the cluster.
2. Use SGI Management Center to start the nodes in the cluster.
   a. Log in as `root`.
   b. Start the SGI Management Center with the following command:

      ```
      # mgrclient
      ```

      The default username/password for SGI Management Center is `root`/`root`.
   c. Within the Management GUI, select the nodes to start, right-click, and select **Power** > **On**.

   MarkLogic is configured to start once the servers have booted.
3. Use the web browser on the head node to log into the MarkLogic Administration web interface:

   http://r01n01:8001

   The default username/password for the MarkLogice interface is `admin`/`admin`.
4. Verify that the cluster powered on correctly and that all slave nodes joined the cluster.

Re-Imaging the Server Nodes

In SGI Management Center, there is a compute image for the MarkLogic compute nodes:

```
Compute-MarkLogic
```

Re-provision the nodes with the compute images as needed. Doing so will not impact the MarkLogic database on the system.

MarkLogic Usage Notes

This chapter describes a couple of usage notes regarding the MarkLogic software:

MarkLogic License Installation

The SGI DataRaptor Appliance with MarkLogic Database comes preset with a permanent license key for the basic MarkLogic sofware. If you have purchased add-on features and do not yet have a license for those options, contact your SGI Sales representative.

MarkLogic Validation Database

The SGI DataRaptor with MarkLogic Database appliance is factory-integrated with a small validation database that consists of one forest per host. The database name is `sgidb` and is provided as a cluster validation. Depending on your application requirements, you may need to create your own validation database. The validation database can be removed at any time but, once removed, it cannot be recreated without support from SGI.

To remove the `sgidb` validation database, run the following commands on node `r01n01` as `root`:

```
# curl --digest -u admin:admin http://localhost:8001/sgi/xqy/delete-xcc.xqy
# curl --digest -u admin:admin http://localhost:8001/sgi/xqy/delete-http.xqy
# curl --digest -u admin:admin http://localhost:8001/sgi/xqy/detach-forest.xqy
# curl --digest -u admin:admin
    http://localhost:8001/sgi/xqy/remove-replica-forest.xqy
# curl --digest -u admin:admin http://localhost:8001/sgi/xqy/delete-forest.xqy
```

```
# curl --digest -u admin:admin http://localhost:8001/sgi/xqy/delete-db.xqy
# curl --digest -u admin:admin http://localhost:8001/sgi/xqy/restart.xqy
```

Verify that the following events occurred:

- The `sgidb-http:10000` and `sgidb-xcc:10001` services have been removed.
- Database `sgidb` has been removed from the Database section.
- The `sgidb` forests have been removed from the Forest section.

Additional Information on SGI Knowledgebase

You can find the customary SGI customer manuals on the SGI Technical Publications Library (See "Related Publications" on page viii.). You can find additional information about the SGI DataRaptor Appliance with MarkLogic Database on SGI Knowledgebase. This knowledgebase is a repository of support information including troubleshooting guides, how-tos, start-heres, and technical solution documents. The following are several usage notes about SGI Knowledgebase:

- SGI Knowledgebase requires an SGI Supportfolio user id and password. If you do not have such validation, go to https://support.sgi.com/.
- The URL for SGI Knowledgebase is https://support.sgi.com/Knowledgebase.
- SGI Knowledgebase has a conventional web search interface. You can find the desired documents by entering some combination of the following terms:
  - `dataraptor`
  - `marklogic`
  - `start here`