

Partitioning Guide for Systems That Contain Cx-bricks

007-4426-001

CONTRIBUTORS

Written by Nancy Heller

Illustrated by Nancy Heller and Dan Young

Production by Karen Jacobson

Engineering contributions by Russ Anderson, Dick Brownell, and Paul Kinyon.

Cover Design By Sarah Bolles, Sarah Bolles Design, and Dany Galgani, SGI Technical Publications.

COPYRIGHT

© 2003, Silicon Graphics, Inc. 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 Silicon Graphics, Inc.

LIMITED RIGHTS LEGEND

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 Silicon Graphics, Inc., 1600 Amphitheatre Pkwy 2E, Mountain View, CA 94043-1351.

TRADEMARKS AND ATTRIBUTIONS

Silicon Graphics, SGI, the SGI logo, InfiniteReality, IRIX, Onyx, and Origin are registered trademarks, and NUMAflex, NUMAlink, and SGIconsole are trademarks, of Silicon Graphics, Inc., in the United States and/or other countries worldwide.

Record of Revision

Version	Description
001	April 2003 Initial version

Contents

Figures	vii
Tables	ix
About This Guide	xi
Related Publications	xi
Cx-brick Addendum	xii
SGI Origin 3000 Series Owner's Guide	xii
SGI Onyx 3000 Series Graphics System Hardware Owner's Guide	xii
IRIX Admin: System Configuration and Operation	xii
Other SGI Documentation	xiii
Online Man Pages	xiii
IRIX Release Notes	xiv
Product Support	xiv
Conventions	xv
Reader Comments	xvi
1. Partitioning	1
Requirements	3
Configurations	4
Creating Partitions	8
part Command	9
mkpart Command	11
setpart Command	13
partmgr Command	14
2. Partition Examples	17
Two 1-brick Partitions	18
Two 2-brick Partitions	20
One 3-brick Partition and One 1-brick Partition	22

Figures

Figure 1-1	Partitioned System	3
Figure 1-2	Brick Placement for a System That Contains Four 128-processor Partitions	5
Figure 2-1	Example of Two 1-brick Partitions	18
Figure 2-2	Brick Placement for Two 1-brick Partitions	19
Figure 2-3	Example of Two 2-brick Partitions	20
Figure 2-4	Brick Placement for Two 2-brick Partitions	21
Figure 2-5	Example of One 3-brick Partition and One 1-brick Partition	22
Figure 2-6	Brick Placement for One 3-brick Partition and One 1-brick Partition.	23
Figure 2-7	Example of One 2-brick Partition and Two 1-brick Partitions.	24
Figure 2-8	Brick Placement for One 2-brick Partition and Two 1-brick Partitions	25
Figure 2-9	Example of Four 1-brick Partitions	26
Figure 2-10	Brick Placement Four 1-brick Partitions	27
Figure 2-11	Example of Three 2-brick Partitions	28
Figure 2-12	Brick Placement for Three 2-brick Partitions	29
Figure 2-13	Example of Four 2-brick Partitions	30
Figure 2-14	Brick Placement for Four 2-brick Partitions	31
Figure 2-15	Example of One 6-brick Partition and One 2-brick Partition	32
Figure 2-16	Brick Placement for One 6-brick Partition and One 2-brick Partition.	33
Figure 2-17	Example of Two 4-brick Partitions	34
Figure 2-18	Brick Placement for Two 4-brick Partitions	35
Figure 2-19	Example of One 4-brick Partition and Two 2-brick Partitions.	36
Figure 2-20	Brick Placement for One 4-brick Partition and Two 2-brick Partitions	37
Figure 2-21	Example of One 4-brick Partition and Four 2-brick Partitions.	38
Figure 2-22	Brick Placement for One 4-brick Partition and Four 2-brick Partitions	39

Figure 2-23	Example of Two 8-brick Partitions 40
Figure 2-24	Brick Placement for Two 8-brick Partitions 41
Figure 2-25	Example of Four 4-brick Partitions 42
Figure 2-26	Brick Placement for Four 4-brick Partitions 43
Figure 2-27	Example of One 14-brick Partition and One 2-brick Partition 44
Figure 2-28	Brick Placement for One 14-brick Partition and One 2-brick Partition 45
Figure 2-29	Example of Eight 2-brick Partitions 46
Figure 2-30	Brick Placement for Eight 2-brick Partitions 47
Figure 2-31	Example of One 8-brick Partition and Two 4-brick Partitions 48
Figure 2-32	Brick Placement for One 8-brick Partition and Two 4-brick Partitions	49
Figure 2-33	Example of One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions 50
Figure 2-34	Brick Placement for One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions 51
Figure 2-35	Example of One 8-brick Partition and Four 2-brick Partitions 52
Figure 2-36	Brick Placement for One 8-brick Partition and Four 2-brick Partitions 53
Figure 2-37	Example of Two 16-brick Partitions 54
Figure 2-38	Brick Placement for Two 16-brick Partitions 55
Figure 2-39	Example of Four 8-brick Partitions 56
Figure 2-40	Brick Placement for Four 8-brick Partitions 57
Figure 2-41	Example of Eight 4-brick Partitions 58
Figure 2-42	Brick Placement for Eight 4-brick Partitions 59
Figure 2-43	Example of One 16-brick Partition and Two 8-brick Partitions 60
Figure 2-44	Brick Placement for One 16-brick Partition and Two 8-brick Partitions 61
Figure 2-45	Example of One 16-brick Partition, One 8-brick Partition, and Two 4-brick Partitions 62
Figure 2-46	Brick Placement for One 16-brick Partition, One 8-brick Partition, and Two 4-brick Partitions 63
Figure 2-47	Example of One 16-brick Partition, One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions 64
Figure 2-48	Brick Placement for One 16-brick Partition, One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions 65

Tables

Table 1-1	Valid System Partitions	6
Table 1-2	mkpart Command Options	11
Table 1-3	partmgr Command Options	15

About This Guide

This guide documents partitioning for systems that contain Cx-bricks (for example, SGI Origin 3900 systems and SGI Onyx 3900 graphics systems). Specifically, it includes the following information:

- Defines partitioning and its benefits.
- Lists the requirements of a partition.
- Lists supported configurations and provides logical and physical drawings of the configurations.
- Describes how to create partitions.

Related Publications

This section discusses SGI documents that are relevant to partitioning a system that contains Cx-bricks, as follows:

- *Cx-brick Addendum*
- *SGI Origin 3000 Series Owner's Guide*
- *SGI Onyx 3000 Series Graphics System Hardware Owner's Guide*
- *IRIX Admin: System Configuration and Operation*
- Other SGI documentation
- Online man pages
- IRIX release notes

Cx-brick Addendum

This addendum (P/N 007-4572-*xxx*) describes the functions, features, physical components, technical specifications, and system configurations for the Cx-brick.

SGI Origin 3000 Series Owner's Guide

This guide (P/N 007-4240-*xxx*) describes the components that compose the SGI Origin 3000 server series. It also provides information about the following procedures:

- Connecting a console to the server system
- Adding or replacing PCI cards and storage disk drives
- Powering on and powering off the server system
- Using the L1 and L2 controllers (and controller commands) to monitor and control the server system

SGI Onyx 3000 Series Graphics System Hardware Owner's Guide

This guide (P/N 007-4264-*xxx*) describes the components that compose an SGI Onyx 3000 rackmount graphics system. It also provides information on using and administering this graphics system.

IRIX Admin: System Configuration and Operation

This guide (P/N 007-2859-*xxx*) lists general system administration practices and describes system administration tasks (for example, partitioning).

Other SGI Documentation

You can obtain SGI documentation in the following ways:

- See the SGI Technical Publications Library at <http://docs.sgi.com>. Various formats are available. This library contains the most recent and most comprehensive set of online books, release notes, man pages, and other information.
- If it is installed on your SGI system, you can use InfoSearch, an online tool that provides a more limited set of online books, release notes, and man pages. With an IRIX system, select **Help** from the Toolchest, and then select **InfoSearch**. Or you can type **infosearch** on a command line.

Online Man Pages

Your system includes a set of IRIX online manual pages, which are formatted in the standard UNIX “man page” style. These man pages are located on the internal system disk (or CD-ROM) and are displayed using the man command. For example, to display the man page for the `Add_disk` command, enter the following command at a shell prompt:

```
man Add_disk
```

Man pages document important system configuration files and commands. References in SGI documentation to these pages include the name of the command and the section number in which the command is located. For example, “`Add_disk(1)`” refers to the `Add_disk` command and indicates that it is located in section 1 of the IRIX reference. For additional information about displaying reference pages using the man command, see `man(1)`.

The `apropos` command locates man pages based on keywords. For example, to display a list of man pages that describe disks, enter the following command at a shell prompt:

```
apropos disk
```

For information about setting up and using `apropos`, see the `apropos(1)` and `makewhatis(1M)` man pages.

IRIX Release Notes

You can view the release notes for various SGI products and software subsystems by using one of two utilities:

`relnotes` Text-based viewer for online release notes.

`grelnotes` Graphics viewer for online release notes.

To see a list of available release notes, type the following command at a shell prompt:

`relnotes`

For more information, see the `relnotes(1)` and `grelnotes(1)` man pages.

Product Support

SGI provides a comprehensive product support and maintenance program for its products, as follows:

- If you are in North America and want support for your SGI-supported products, contact the Technical Assistance Center at +1 800 800 4SGI or your authorized service provider.
- If you are outside North America, contact the SGI subsidiary or authorized distributor in your country.

Conventions

The following conventions are used throughout this document:

Convention	Meaning
<code>command</code>	This fixed-space font denotes literal items such as commands, files, routines, path names, signals, messages, and programming language structures.
<i>variable</i>	Italic typeface denotes variable entries and words or concepts being defined.
user input	This bold, fixed-space font denotes literal items that the user enters in interactive sessions. Output is shown in nonbold, fixed-space font.
[]	Brackets enclose optional portions of a command or directive line.
...	Ellipses indicate that a preceding element can be repeated.
<code>manpage(x)</code>	Man page section identifiers appear in parentheses after man page names.
GUI element	This font denotes the names of graphical user interface (GUI) elements such as windows, screens, dialog boxes, menus, toolbars, icons, buttons, boxes, fields, and lists.

Reader Comments

If you have comments about the technical accuracy, content, or organization of this document, please 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
- Use the Feedback option on the Technical Publications Library Web page:
<http://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:
Technical Publications
SGI
1600 Amphitheatre Pkwy, M/S 535
Mountain View, California 94043-1351
- Send a fax to the attention of “Technical Publications” at +1 650 932 0801.

SGI values your comments and will respond to them promptly.

Partitioning

SGI systems that use the NUMAflex design (such as SGI Origin 3000 and SGI Onyx 3000 series systems) and run the IRIX 6.5 or later release can be partitioned. Partitioning divides a multi-node distributed shared memory (DSM) system into smaller systems without the need to physically recable the system. Each partition, which is logically separated from other partitions, runs its own operating system kernel and behaves as a standalone system. Each partition has its own console, root filesystem (system I- or IX-brick), and IP (internet protocol) address.

Partitioning provides the following benefits:

- High availability. Any SGI NUMAflex server can be partitioned to appear as a cluster of systems. Unlike a cluster of systems, however, it benefits from the low-latency and high-bandwidth NUMAlink communication between partitions. NUMAlink and partitioning provide both high availability and high performance.
- Job resilience. Jobs that run within a partition can use checkpointing to save job information, and each partition can be checkpointed individually as a single memory and application space. Checkpointed jobs can be restarted after the partition has been repaired, or they can be restarted in another partition. Load-balancing software (such as Load Sharing Facility [LSF] from Platform Computing) can manage workloads between partitions in a system the same as in a cluster. The IRIX Advanced Cluster Environment provides a checkpoint and restart mechanism for jobs that run on more than one partition.
- No single point of failure. A single shared-memory system can be divided into separate partitions, eliminating the possibility of a single point of failure for the entire system.

- Multiple environments. Because each partition in the system is running a separate operating system kernel, different IRIX kernels can be running in a single system at the same time, allowing the flexibility to manage all types of users and jobs. You can use one partition for development or experimental work and another partition for production work.
- Fault isolation. Failures in one partition will not affect the other partitions. Upon an unrecoverable hardware- or software-related failure, the affected partition immediately shuts down, limiting the propagation of data corruption within the partition. Memory protection hardware within all partitions prevents bad packets from corrupting memory.

Figure 1-1 provides an example of a system that is divided into two partitions. Both partitions can be booted, powered off, powered on, and rebooted without affecting the normal operation of the other partition.

The NUMALink interconnect is used for communication within partitions and between partitions. The NUMALink interconnect uses memory-to-memory copy routines, so inter-partition communication is a high-performance alternative to I/O clustering. Each NUMALink cable can transfer data at a rate of 3.2 GB/s.

For communication between partitions, you can use either of the following mechanisms:

- Network interface driver using sockets (such as TCP/IP or UDP)
- Shared-memory interface

Note: The TCP/IP interface supports both NFS and CXFS traffic between partitions.

Requirements

Each partition must have the infrastructure to run as a standalone system. This infrastructure includes the following components:

- At least one Cx-brick for system sizes of four Cx-bricks or less, or multiples of two Cx-bricks for system sizes greater than four Cx-bricks

Note: The Cx-bricks do not have to be fully populated with processors.

- One I- or IX-brick
- One `root` disk that contains the operating system image
- One console connection

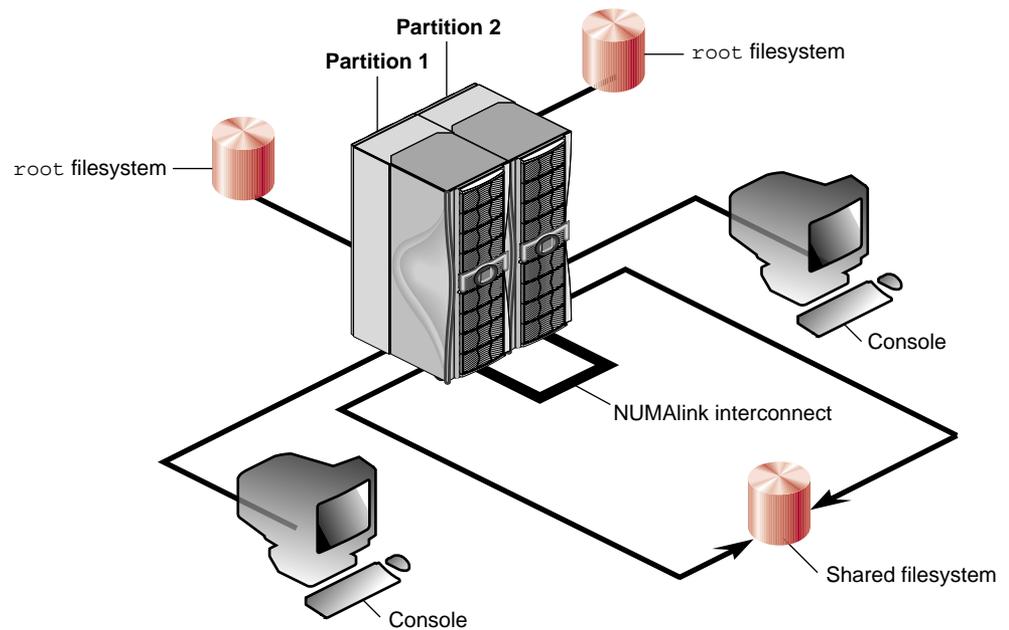


Figure 1-1 Partitioned System

Configurations

You can configure your system as a single system, or you can divide it into partitions that operate as separate systems. For example, you can configure an Origin 3900 system that has 512 processors as a single 512-processor system, or you can divide it into as many as eight separate systems. The size of a partition can be configured to match your job size.

Reconfiguring partitions is performed through software only and does not require any cabling changes. For example, you can configure your system with multiple partitions during the day and then reconfigure the system as a single partition to run a large application at night. You can perform all of this configuring from a remote location. For more information about reconfiguring, see “Creating Partitions” on page 8.

The following configuration rules apply to partitions:

- Each partition must have a unique ID number between 1 and 63, inclusively. An ID number of 0 indicates that the system is not partitioned.
- The console of the lowest-numbered Cx-brick in a partition becomes the console for that partition. (The original console remains the console for the partition that contains it.)
- Each partition has its own set of PROM environment variables: `ConsolePath`, `OSLoadPartition`, `SystemPartition`, `netaddr`, and `root`.
- A partition must be contiguous in topology. The route between any two processors in the same partition must be contained within that partition, and not through any other partition. If the bricks in a partition are not contiguous, the system will not boot.
- An I/O brick belongs to the partition to which the attached Cx-brick belongs. If an I/O brick is attached to two Cx-bricks, both Cx-bricks must be in the same partition. I/O bricks cannot be shared by two partitions.

Note: Peripherals, such as dual-ported disks, can be shared the same way that two nodes in a cluster can share peripherals.

- The G-brick must follow the same partitioning rules as the I/O bricks. Both graphics pipes on a G-brick must be connected to the same partition.
- For Origin 3900 systems that contain more than 18 Cx-bricks, specific bricks must reside in each partition. In these configurations, the Cx-bricks reside in three or four

compute racks. Partitions that consist of more than two Cx-bricks must be “striped” across the compute racks. For example, in a 512-processor system that is partitioned as four 128-processor partitions, the partitions consist of the following Cx-bricks (see Figure 1-2):

- Partition 1 consists of Cx-bricks 1 through 8
- Partition 2 consists of Cx-bricks 9 through 16
- Partition 3 consists of Cx-bricks 17 through 24
- Partition 4 consists of Cx-bricks 25 through 32

Front view of racks

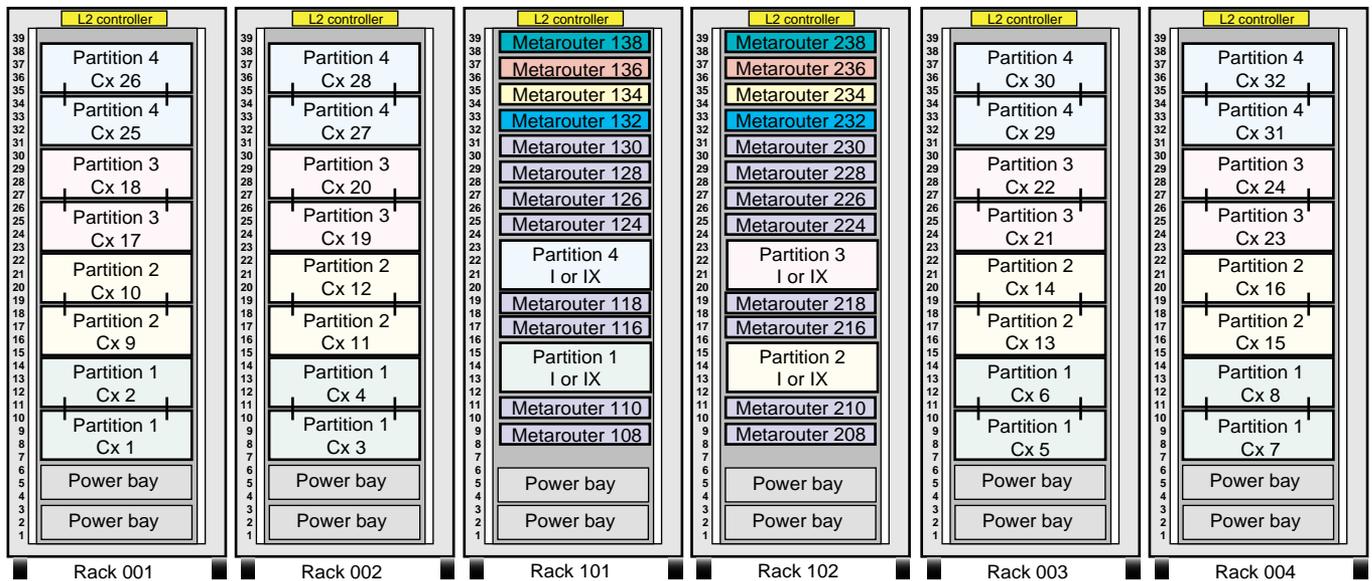


Figure 1-2 Brick Placement for a System That Contains Four 128-processor Partitions

Table 1-1 lists the partition configurations that are available to the Origin 3900 system. The marketing codes for systems that contain eight Cx-bricks or less are also valid for Onyx 3900 graphics systems. Illustrations of the partitioned systems are provided in Chapter 2.

Table 1-1 Valid System Partitions

Marketing Code	Number of Cx-bricks in System	Number of Partitions	Number of Cx-bricks in Each Partition	Maximum Number of Processors in Partition	Illustrations
Partition-O3900-C020	2	2	1	16	Figure 2-1 on page 18 and Figure 2-2 on page 19
Partition-O3900-D030	4	2	2	32	Figure 2-3 on page 20 and Figure 2-4 on page 21
Partition-O3900-D031	4	4	1	16	Figure 2-9 on page 26 and Figure 2-10 on page 27
Partition-O3900-D034	4	3	1 partition with 2 Cx-bricks 2 partitions with 1 Cx-brick	32 16	Figure 2-7 on page 24 and Figure 2-8 on page 25
Partition-O3900-D037	4	2	1 partition with 3 Cx-bricks 1 partition with 1 Cx-brick	48 16	Figure 2-5 on page 22 and Figure 2-6 on page 23
Partition-O3900-E050	8	2	4	64	Figure 2-17 on page 34 and Figure 2-18 on page 35
Partition-O3900-E051	8	4	2	32	Figure 2-13 on page 30 and Figure 2-14 on page 31
Partition-O3900-E054	8	3	1 partition with 4 Cx-bricks 2 partitions with 2 Cx-bricks	64 32	Figure 2-19 on page 36 and Figure 2-20 on page 37
Partition-O3900-E057	8	2	1 partition with 6 Cx-bricks 1 partition with 2 Cx-bricks	96 32	Figure 2-15 on page 32 and Figure 2-16 on page 33
Partition-O3900-F070	16	2	8	128	Figure 2-23 on page 40 and Figure 2-24 on page 41
Partition-O3900-F071	16	4	4	64	Figure 2-25 on page 42 and Figure 2-26 on page 43
Partition-O3900-F072	16	8	2	32	Figure 2-29 on page 46 and Figure 2-30 on page 47

Table 1-1 Valid System Partitions **(continued)**

Marketing Code	Number of Cx-bricks in System	Number of Partitions	Number of Cx-bricks in Each Partition	Maximum Number of Processors in Partition	Illustrations
Partition-O3900-F076	16	3	1 partition with 8 Cx-bricks 2 partitions with 4 Cx-bricks	128 64	Figure 2-31 on page 48 and Figure 2-32 on page 49
Partition-O3900-F077	16	4	1 partition with 8 Cx-bricks 1 partition with 4 Cx-bricks 2 partitions with 2 Cx-bricks	128 64 32	Figure 2-33 on page 50 and Figure 2-34 on page 51
Partition-O3900-F078	16	5	1 partition with 8 Cx-bricks 4 partitions with 2 Cx-bricks	128 32	Figure 2-35 on page 52 and Figure 2-36 on page 53
Partition-O3900-F079	16	2	1 partition with 14 Cx-bricks 1 partition with 2 Cx-bricks	224 32	Figure 2-27 on page 44 and Figure 2-28 on page 45
Partition-O3900-G090	32	2	16	256	Figure 2-37 on page 54 and Figure 2-38 on page 55
Partition-O3900-G091	32	4	8	128	Figure 2-39 on page 56 and Figure 2-40 on page 57
Partition-O3900-G092	32	8	4	64	Figure 2-41 on page 58 and Figure 2-42 on page 59
Partition-O3900-G095	32	3	1 partition with 16 Cx-bricks 2 partitions with 8 Cx-bricks	256 128	Figure 2-43 on page 60 and Figure 2-44 on page 61
Partition-O3900-G096	32	4	1 partition with 16 Cx-bricks 1 partition with 8 Cx-bricks 2 partitions with 4 Cx-bricks	256 128 64	Figure 2-45 on page 62 and Figure 2-46 on page 63
Partition-O3900-G097	32	5	1 partition with 16 Cx-bricks 1 partition with 8 Cx-bricks 1 partition with 4 Cx-bricks 2 partitions with 2 Cx-bricks	256 128 64 32	Figure 2-47 on page 64 and Figure 2-48 on page 65
Partition-O3900-H105	6	3	2	32	Figure 2-11 on page 28 and Figure 2-12 on page 29
Partition-O3900-H120	12	5	1 partition with 4 Cx-bricks 4 partitions with 2 Cx-bricks	64 32	Figure 2-21 on page 38 and Figure 2-22 on page 39

Creating Partitions

You can set up or reconfigure partitions by using the following commands:

- `part` command from the L2 controller
- `mkpart` partition configuration command from the IRIX operating system
- `setpart` command from the command monitor

You can use the `partmgr` command to invoke a partitioning administration tool that enables you to define and store partition definitions.

Note: The `mkpart` and `partmgr` commands can be run only by the system administrator with superuser privilege. If the system is configured as `CAP_NO_SUPERUSER`, you must have effective capabilities of `CAP_SHUTDOWN`, `CAP_DEVICE_MGT`, and `CAP_SYSINFO_MGT`. See the `capabilities(4)` man page for more information on the capability mechanism that provides fine-grained control over the privileges of a process.

Each partition has its own system console from which you can run these commands, or you can use `SGIconsole` to manage multiple partitions.

part Command

To partition a system via an L2 controller, use the `part` command to set the partition ID for each Cx-brick in a partition. For example, to place the Cx-brick in rack 001, slot U07, in partition 1, enter the following command:

```
L2>r 1 s 7 brick part 1
```

Note: To access the L2 controller, type `Ctrl t`. The L2 prompt (L2>) should appear on the screen.

The following list of commands partitions a 512-processor system into four 128-processor partitions:

```
L2>r 1 s 7,11 brick part 1
L2>r 2 s 7,11 brick part 1
L2>r 3 s 7,11 brick part 1
L2>r 4 s 7,11 brick part 1
L2>r 1 s 15,19 brick part 2
L2>r 2 s 15,19 brick part 2
L2>r 3 s 15,19 brick part 2
L2>r 4 s 15,19 brick part 2
L2>r 1 s 23,27 brick part 3
L2>r 2 s 23,27 brick part 3
L2>r 3 s 23,27 brick part 3
L2>r 4 s 23,27 brick part 3
L2>r 1 s 31,35 brick part 4
L2>r 2 s 31,35 brick part 4
L2>r 3 s 31,35 brick part 4
L2>r 4 s 31,35 brick part 4
```

You can verify that the partition IDs are set correctly by using the `cfg L2 controller` command. This command identifies partitions by placing a number after the brick identifier. For example, `001c07.1` indicates that the Cx-brick in slot 07 of rack 001 is in partition 1, and `001c35.4` indicates that the Cx-brick in slot 35 of rack 001 is in partition 4.

In addition, you can use the `sel L2` command to list the system consoles in a partitioned system. For example, the following output indicates the system consoles for a system that contains two partitions:

```
L2>sel
known system consoles (partitioned)
    Partition 1: 001c07 - L2 detected
    Partition 2: 001c11 - L2 detected
current system console
console input: partition 2, 001c11 console0
console output: any brick partition 2
L2>
```

To enable communication between a partition and the L2 controller, use the `sel p L2` command. This command selects which partition is the console. For example, enter **sel p 1** to select partition 1 as the console. When partition 1 is the console, you can access any brick within this partition via the L2 controller.

Note: To select each partition as the console, you must use multiple windows, one window for each partition. Because each partition acts as a standalone system, these windows enable you to access each partition via the L2 controller.

mkpart Command

The `mkpart` partition configuration command allows you to perform the following tasks:

- List the current partitions.
- Save the current configuration.
- Set up partitions as follows:
 - Split an existing system into multiple partitions.
 - Coalesce partitions into one partition.
 - Split up existing partitions and regroup them differently.

See the `mkpart (1M)` man page for detailed information on this command. Table 1-2 lists the `mkpart` command options.

Table 1-2 `mkpart` Command Options

Option	Description
-p partid	Specifies the partition ID for the new configuration.
-m brick	Indicates the brick number.
-i	Coalesces all currently active partitions into a single system with a partition ID of 0. A partition ID of 0 indicates a system that is not partitioned.
-l	Lists current partitions.
-n	Prints partition ID of this partition.
-h	Prints a usage message.
-F	Forces override of all sanity checks. The command sets up partitioning as given and tries to reboot the system. Caution: Use the -F option only if you fully understand the consequences of using this option incorrectly.

A typical sequence of partitioning commands is as follows:

1. Run `mkpart -l` to list all partitions.
2. Run `mkpart` again to repartition the system.
 - Partition IDs are stored in the L1 controllers.
 - `mkpart` uses a daemon (`mkpd`) to write IDs on remote partition L1 controllers.
3. Reboot all affected partitions if `mkpart` is successful.

During the boot operation, each Cx-brick in a system is assigned a module identifier (ID) and a partition ID that is maintained by the brick's L1 controller. By design, the module ID is unique for each Cx-brick, while all Cx-bricks in a single partition have the same partition ID.

When a partition is booted initially, a dynamic discovery routine identifies new hardware in the system. After the discovery process, a reset fence is created between a hub or router on the other side of the link that has a different partition ID. The dynamic discovery and fence generation for a partition occurs while the other partitions continue operation.

setpart Command

You can also set up partitions by using the `setpart` command from the command monitor. To use the `setpart` command, follow these steps:

1. Reset the system.
2. From the System Maintenance menu, enter option 5 for the command monitor.
3. Run the `setpart` command. (The `setpart -h` command displays all `setpart` options.) The `setpart` command is an interactive command that receives your input, including a partition number and the module identifier of the Cx-bricks that make up the partition.
4. Reset the system again and the system will function as two or more partitions.

To coalesce multiple partitions into one partition, follow these steps:

1. Access the command monitor for all the partitions.
2. Run the `setpart -c` command in any of the partitions. This command clears all partition information and offers to reset all partitions simultaneously. After the reset, the system is configured as a single partition.

partmgr Command

The `partmgr` command invokes a partitioning administration tool for the Origin 3000 series of servers that enables you to define and store partition definitions, including the following:

- Partition ID information for bricks and racks.
- Associated PROM environment variables.

After using the `partmgr` command, you must reboot all the affected partitions in order to change the system to the new partition scheme.

Note: Only a system administrator with superuser privilege can use this command. If the system is configured as `CAP_NO_SUPERUSER`, you must have effective capabilities of `CAP_SHUTDOWN`, `CAP_DEVICE_MGT`, and `CAP_SYSINFO_MGT`. See the `capabilities(4)` man page for more information on the capability mechanism that provides fine-grained control over the privileges of a process.



Caution: If you use the `partmgr` command incorrectly, this command can completely freeze all partitions in the system and prevent the partitions from booting the PROM.

To use the `partmgr` command, you must be sure that the following requirements are met:

- You must install the `oe.sw.partition` software package, after which you must configure a new kernel and boot the system.
- The L1 and L2 firmware must be current.
- You must enable partitioning administration in all L1 controllers using the `partdb enable L1` command.
- The `mkpd` daemon must be running on all partitions for `partmgr` to work. If it is not running, the `partmgr` command may not be able to shut down remote partitions.
- The PROM environment variable `AutoPROMVars` must be set to 1. This enables the PROM to use the environment variables defined in the mode definition.

See the `partmgr (1M)` man page for detailed information on this command. Table 1-3 describes the command options.

Table 1-3 `partmgr` Command Options

Option	Description
-h	Help. Prints a usage message.
-l	Lists brick IDs of all known system bricks to standard out.
-a	Lists all currently defined modes.
-n	Displays the current partition ID.
-v	Displays the environment variables for the current partition.
-c	Displays mode configuration. If no mode is specified (-m modename), the current mode is displayed.
-x modename	Deletes a mode.
-s modename	Sets a mode as the current mode. This may require some partitions to be rebooted.
-i	Same as “-s Unpartitioned.” Initializes the system to the default unpartitioned state.
-d	Defines a new node.
-m modename	Specifies a mode name. This option is used with the -c and -d options.
-u	When you define a new mode, this option specifies an unpartitioned system.
-p part	Specifies a partition ID. A partition ID is a number between 1 and 63, inclusive. The order of the -p, -b, -r, and -e options is significant. See the <code>partmgr (1M)</code> man page.
-b bricks	Specifies brick IDs (example: 001c13).
-r racks	Specifies rack IDs (example: 001).
-e var=value	Specifies an environment variable name/value pair.
-f file	When you define a new mode, this option specifies an input file.
-F	Forces override of some sanity checks. Use this option with caution and only if you fully understand the consequences of using this option incorrectly.

Partition Examples

This chapter contains partitioning examples of the Origin 3900 system, as follows:

- “Two 1-brick Partitions” on page 18
- “Two 2-brick Partitions” on page 20
- “One 3-brick Partition and One 1-brick Partition” on page 22
- “One 2-brick Partition and Two 1-brick Partitions” on page 24
- “Four 1-brick Partitions” on page 26
- “Three 2-brick Partitions” on page 28
- “Four 2-brick Partitions” on page 30
- “One 6-brick Partition and One 2-brick Partition” on page 32
- “Two 4-brick Partitions” on page 34
- “One 4-brick Partition and Two 2-brick Partitions” on page 36
- “One 4-brick Partition and Four 2-brick Partitions” on page 38
- “Two 8-brick Partitions” on page 40
- “Four 4-brick Partitions” on page 42
- “One 14-brick Partition and One 2-brick Partition” on page 44
- “Eight 2-brick Partitions” on page 46
- “One 8-brick Partition and Two 4-brick Partitions” on page 48
- “One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions” on page 50
- “One 8-brick Partition and Four 2-brick Partitions” on page 52
- “Two 16-brick Partitions” on page 54
- “Four 8-brick Partitions” on page 56
- “Eight 4-brick Partitions” on page 58
- “One 16-brick Partition and Two 8-brick Partitions” on page 60
- “One 16-brick Partition, One 8-brick Partition, and Two 4-brick Partitions” on page 62
- “One 16-brick Partition, One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions” on page 64

Two 1-brick Partitions

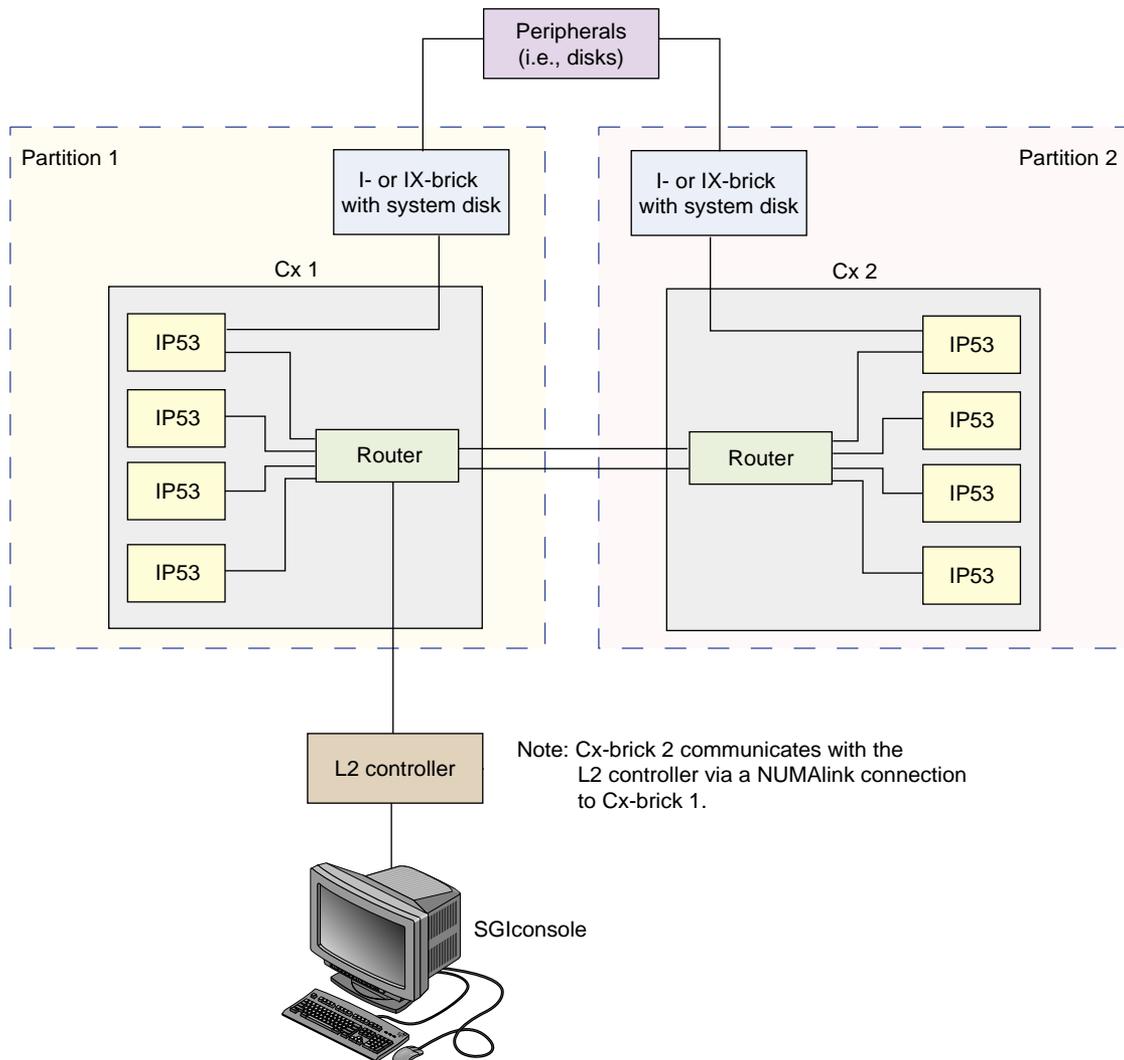


Figure 2-1 Example of Two 1-brick Partitions

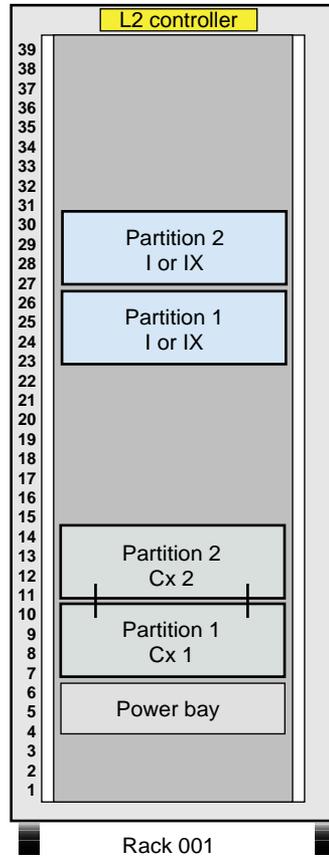


Figure 2-2 Brick Placement for Two 1-brick Partitions

Two 2-brick Partitions

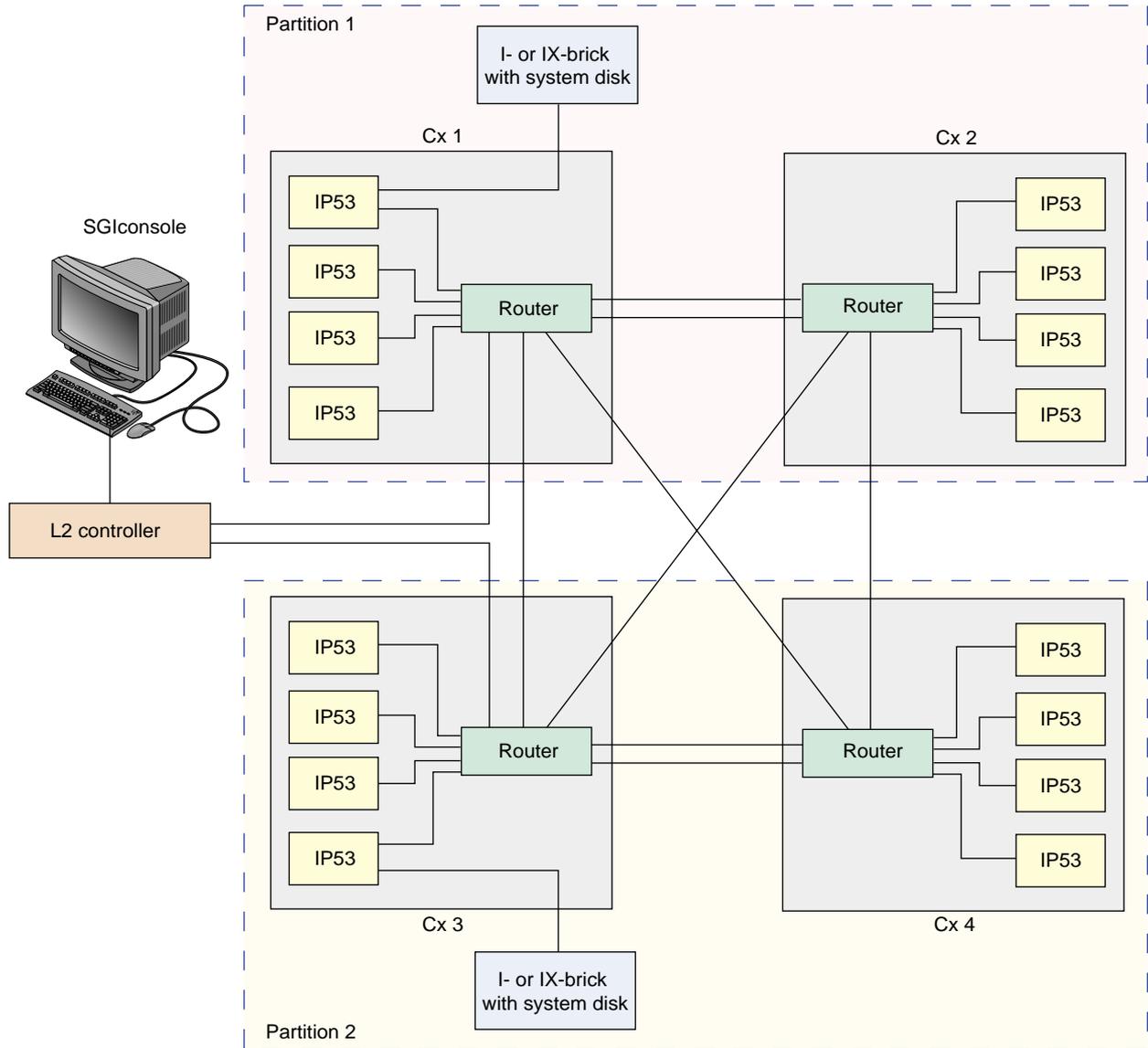


Figure 2-3 Example of Two 2-brick Partitions

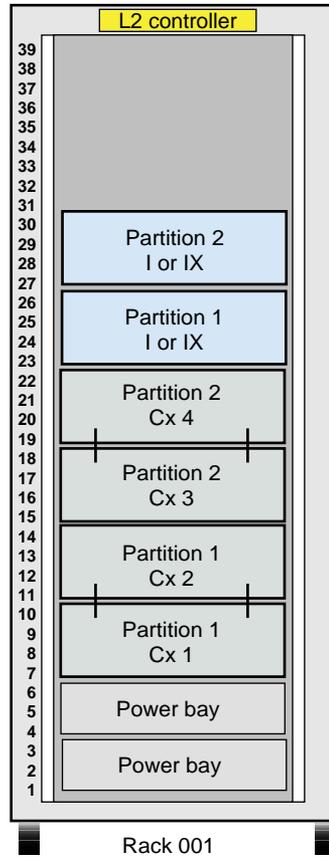


Figure 2-4 Brick Placement for Two 2-brick Partitions

One 3-brick Partition and One 1-brick Partition

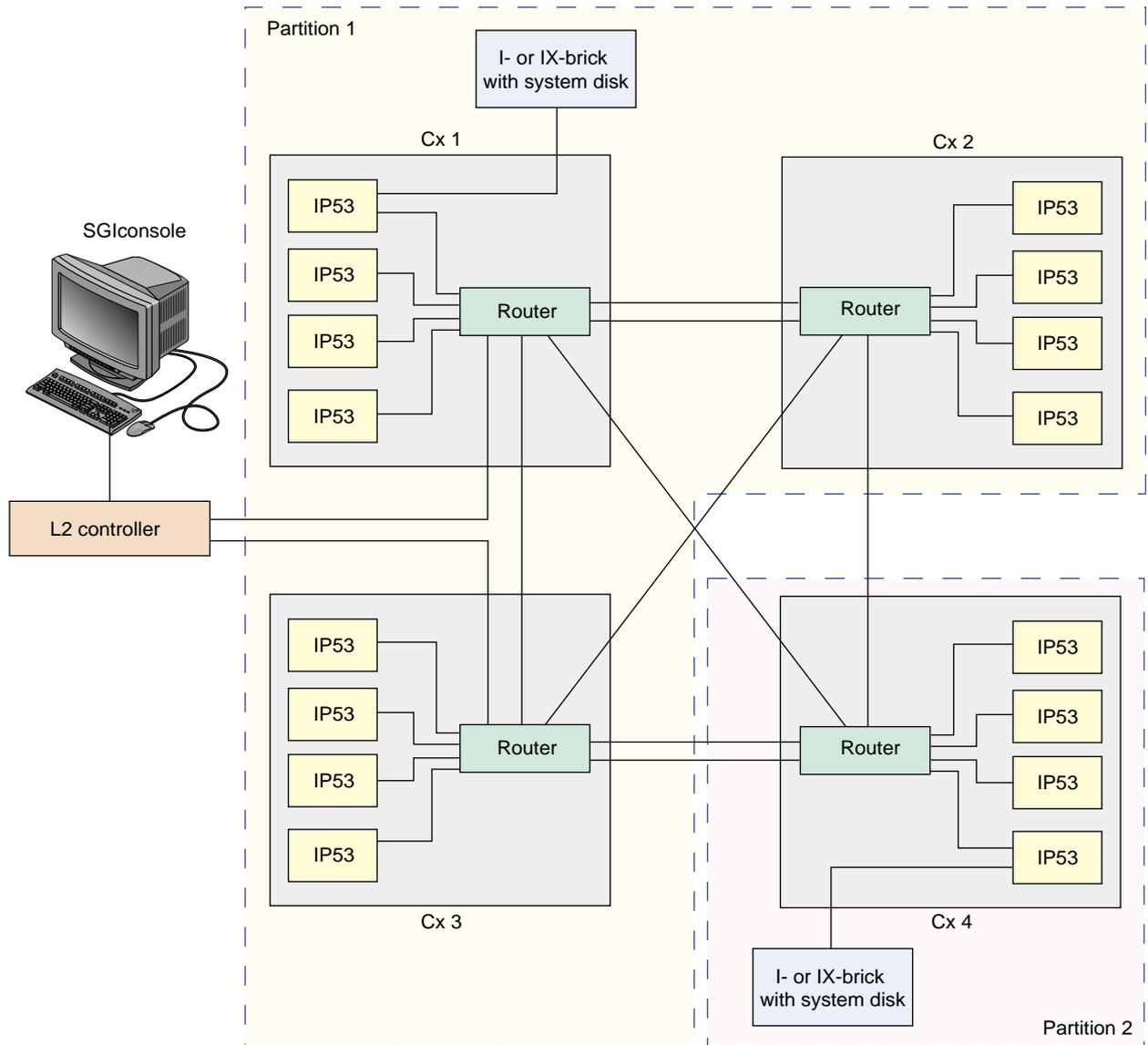


Figure 2-5 Example of One 3-brick Partition and One 1-brick Partition

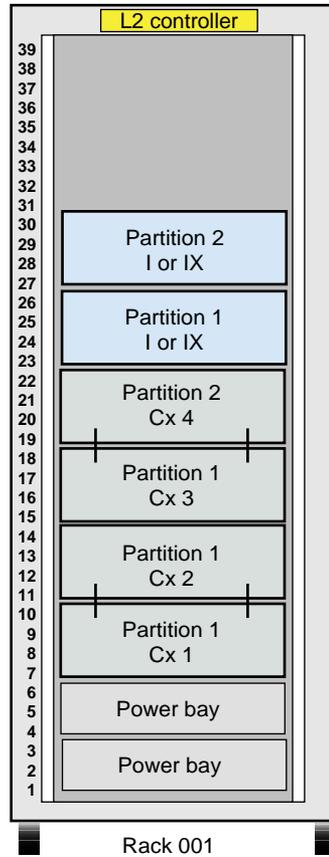


Figure 2-6 Brick Placement for One 3-brick Partition and One 1-brick Partition

One 2-brick Partition and Two 1-brick Partitions

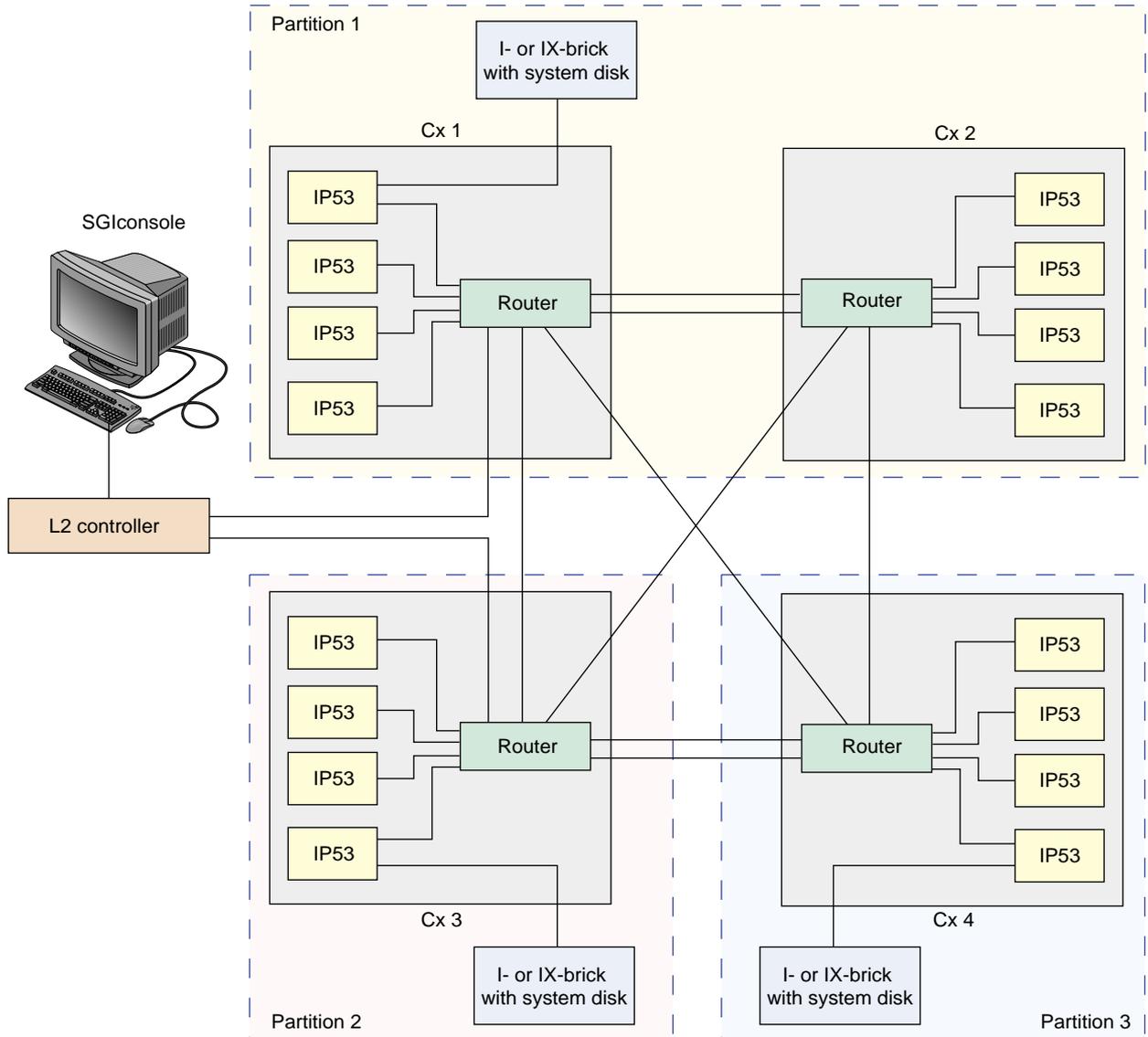


Figure 2-7 Example of One 2-brick Partition and Two 1-brick Partitions

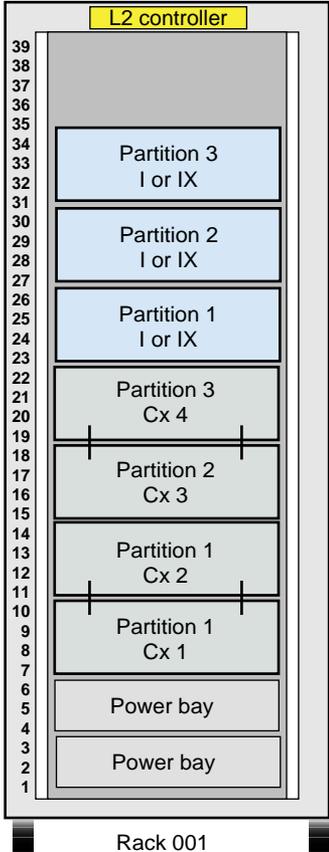


Figure 2-8 Brick Placement for One 2-brick Partition and Two 1-brick Partitions

Four 1-brick Partitions

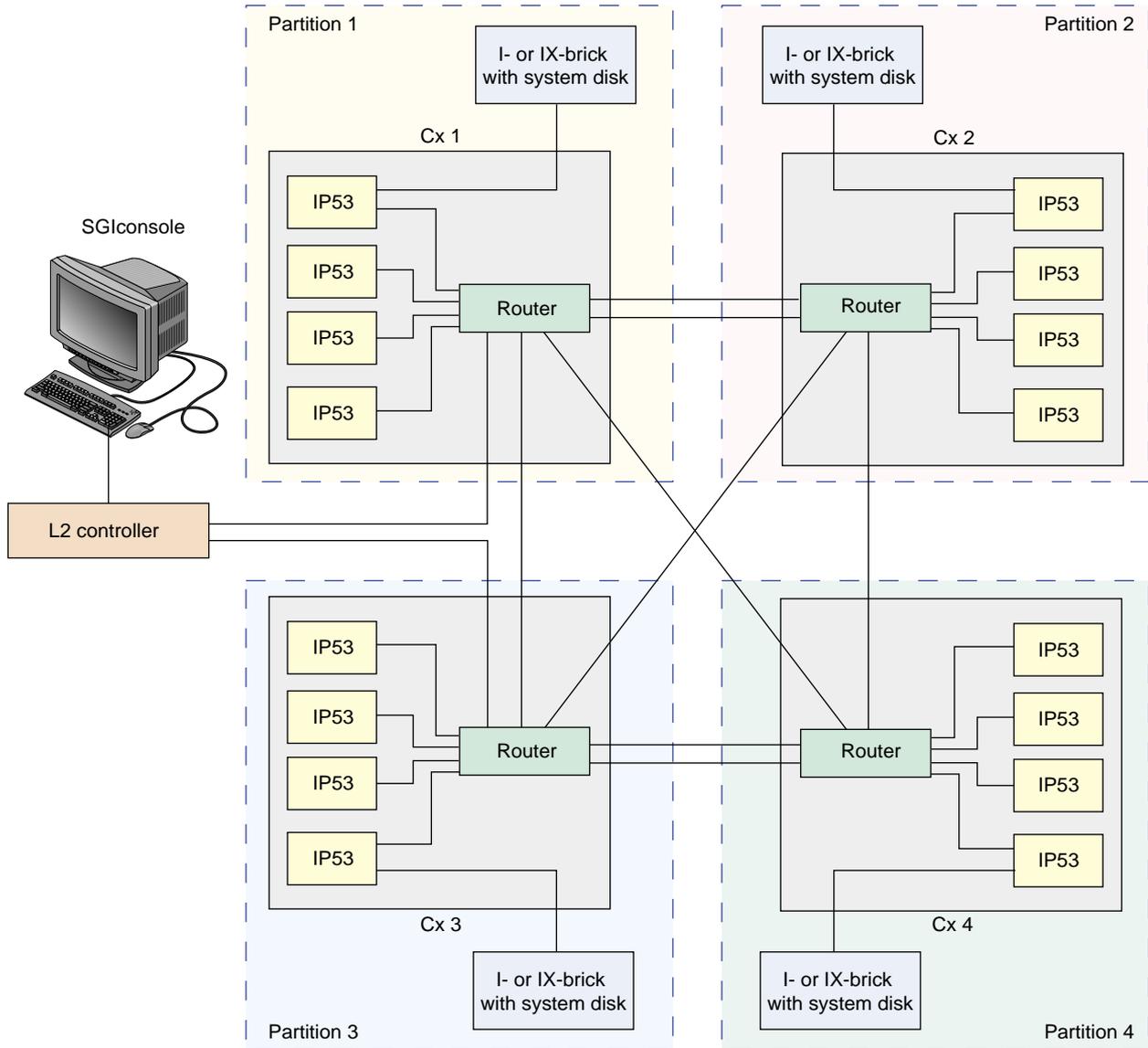


Figure 2-9 Example of Four 1-brick Partitions

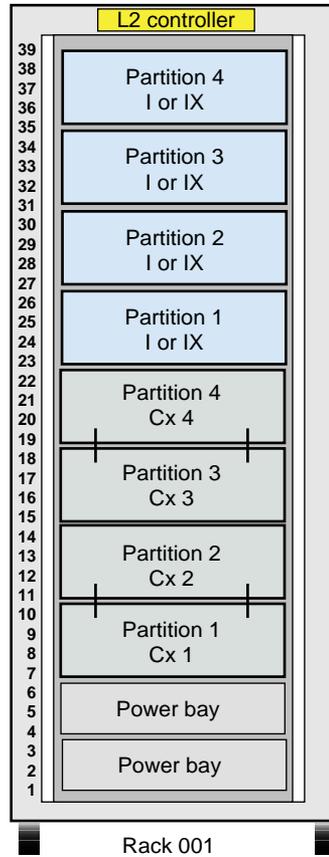


Figure 2-10 Brick Placement Four 1-brick Partitions

Three 2-brick Partitions

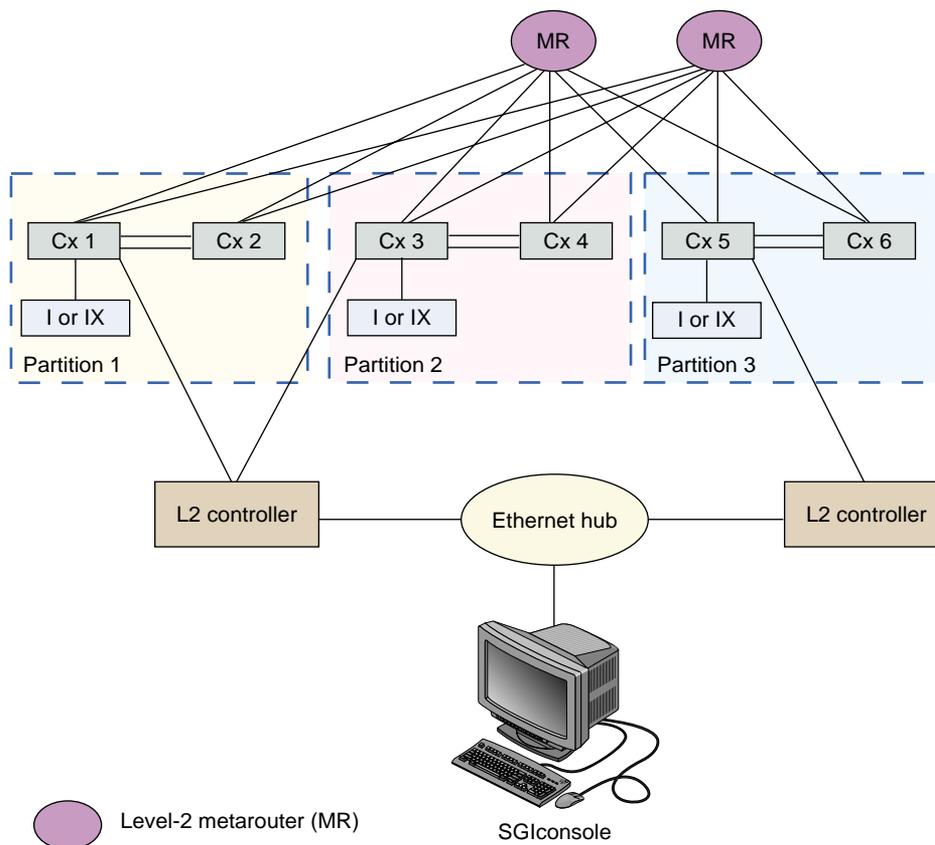


Figure 2-11 Example of Three 2-brick Partitions

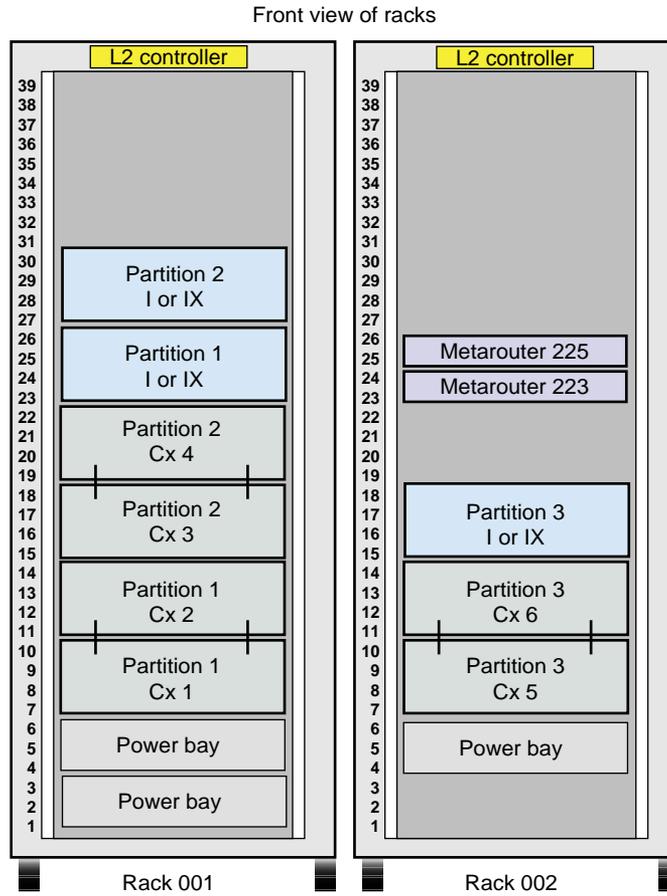


Figure 2-12 Brick Placement for Three 2-brick Partitions

Four 2-brick Partitions

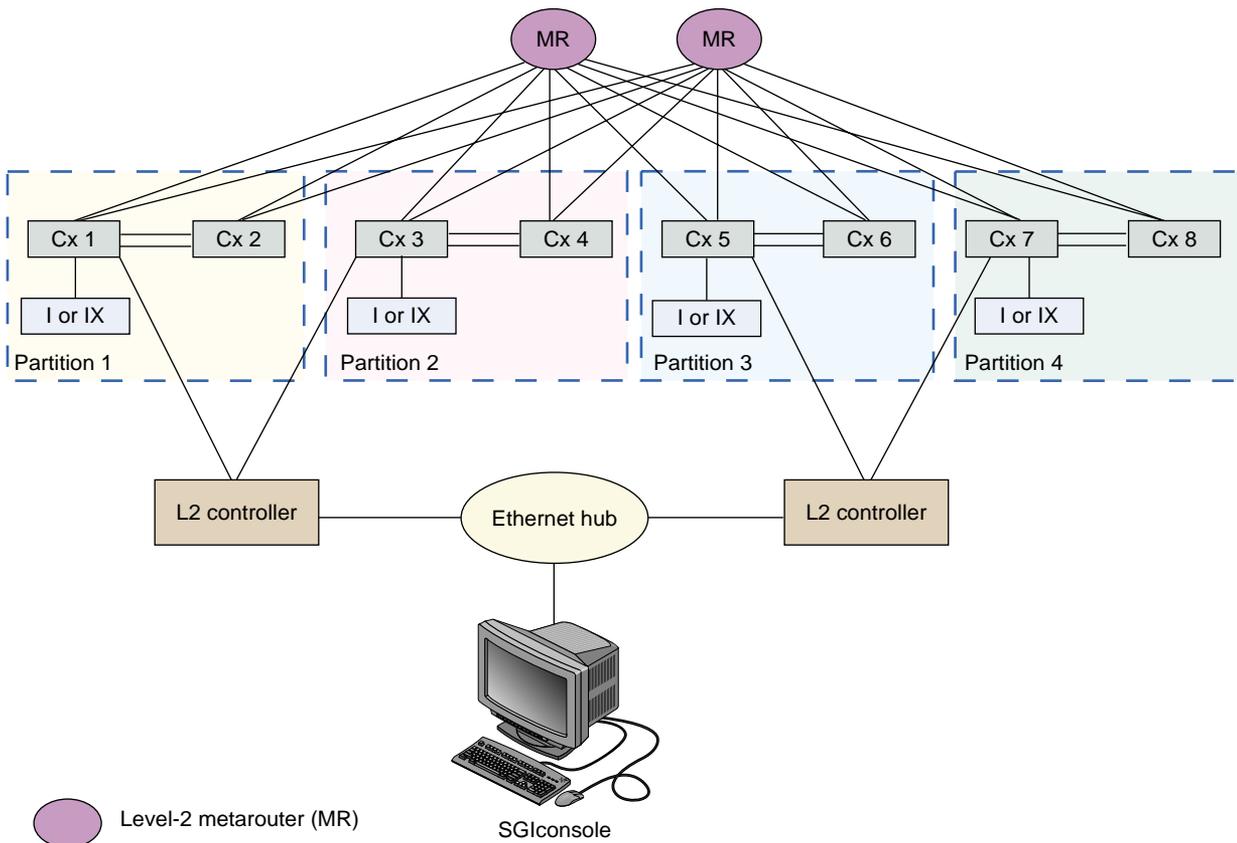


Figure 2-13 Example of Four 2-brick Partitions

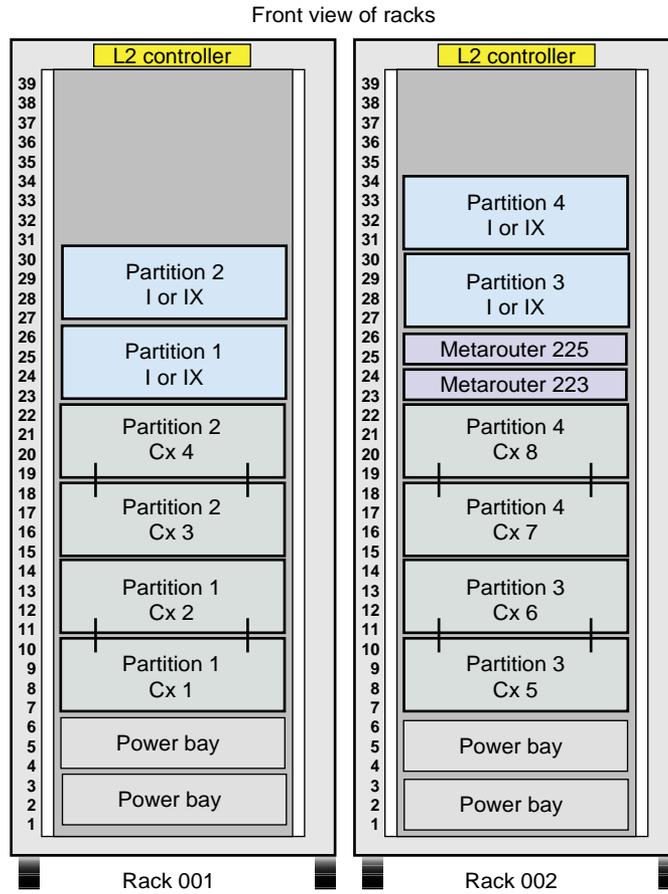


Figure 2-14 Brick Placement for Four 2-brick Partitions

One 6-brick Partition and One 2-brick Partition

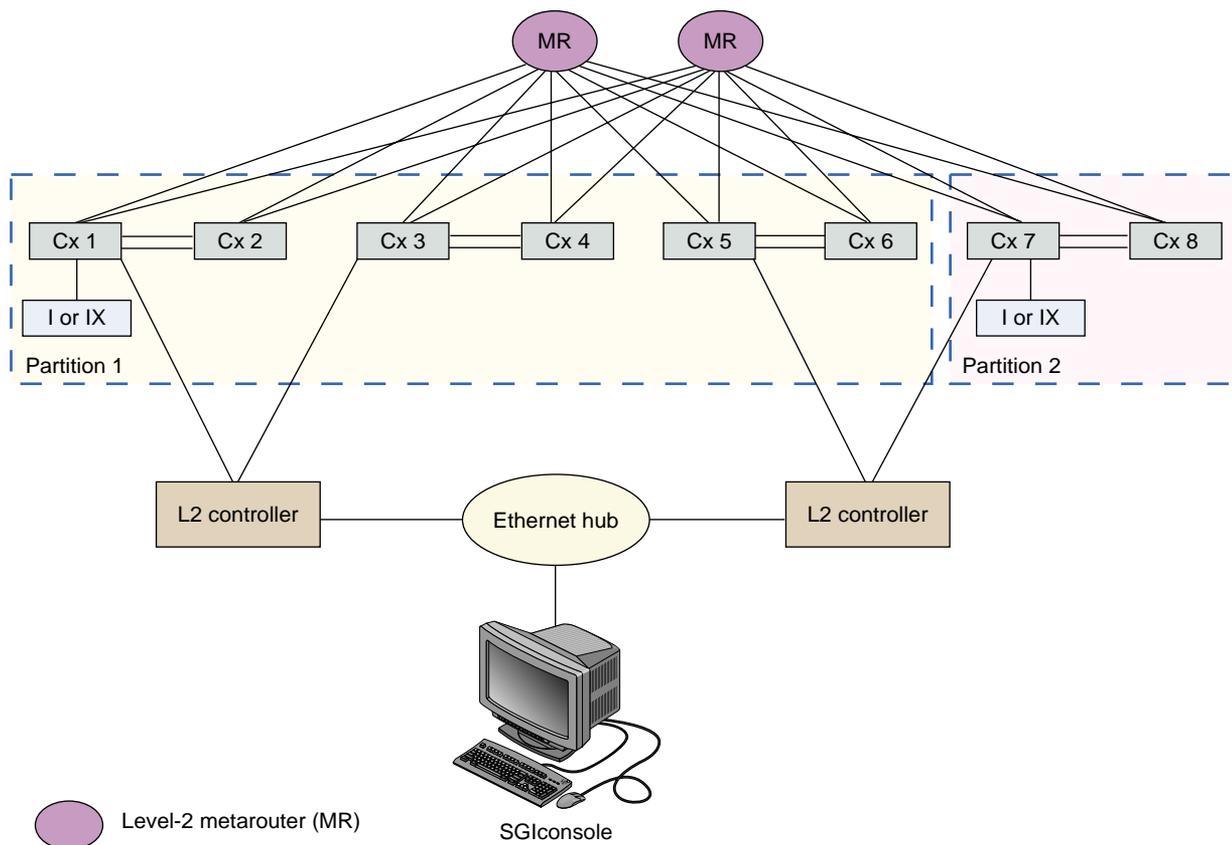


Figure 2-15 Example of One 6-brick Partition and One 2-brick Partition

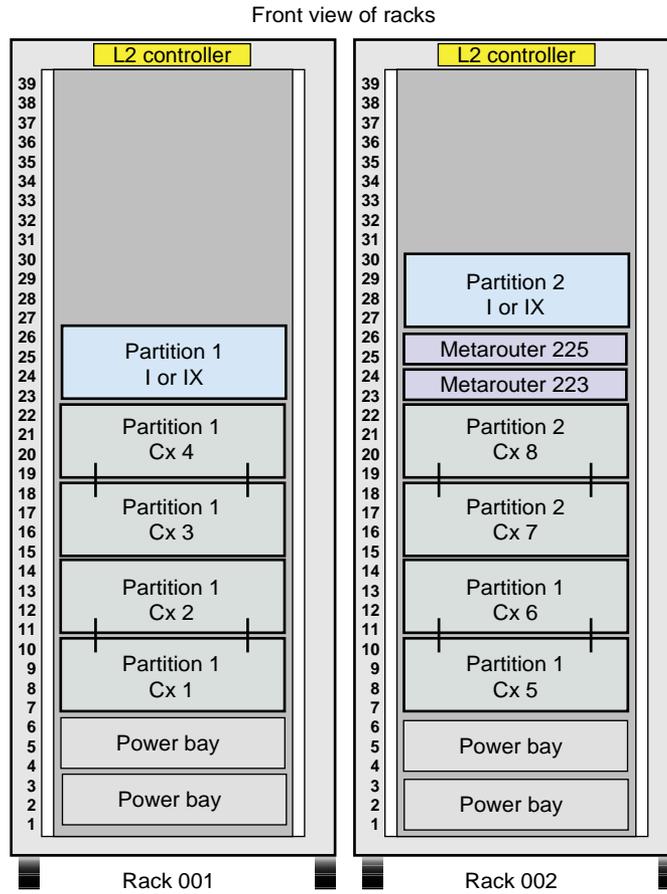


Figure 2-16 Brick Placement for One 6-brick Partition and One 2-brick Partition

Two 4-brick Partitions

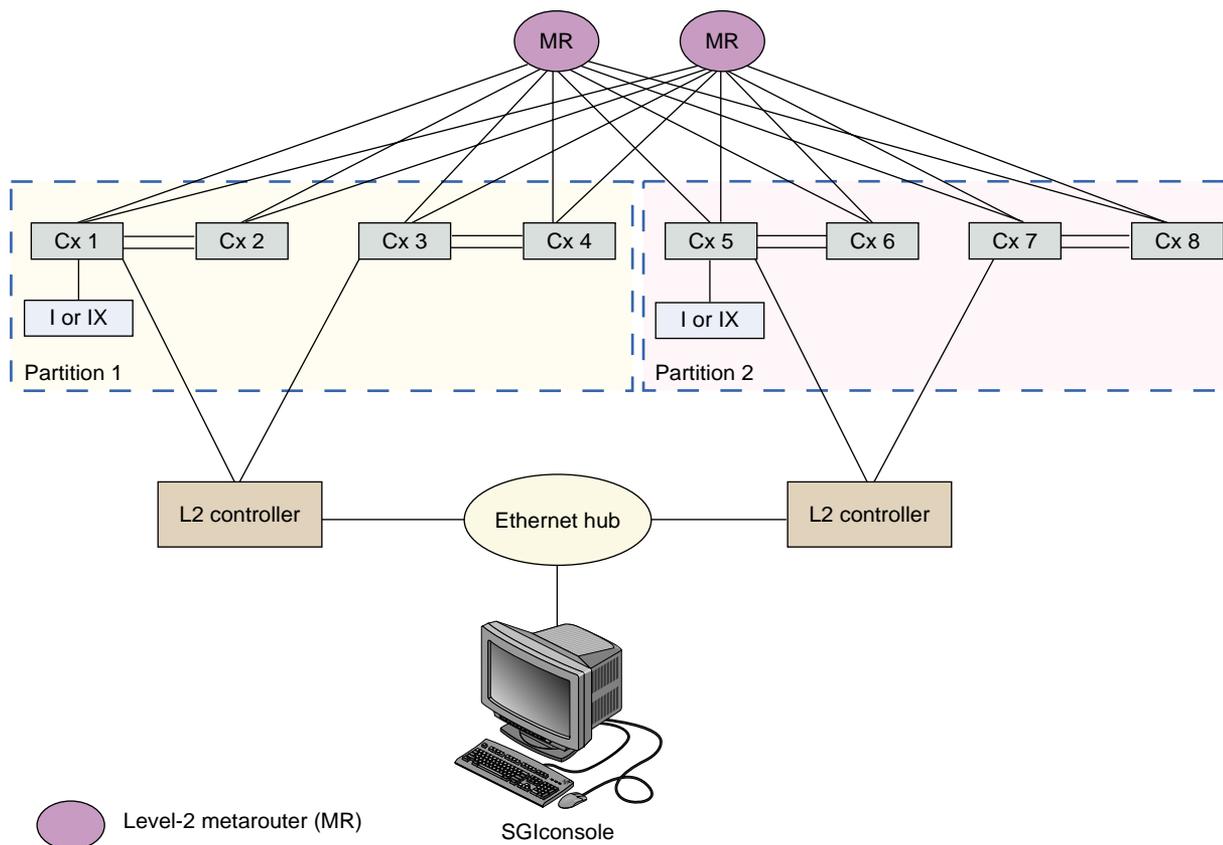


Figure 2-17 Example of Two 4-brick Partitions

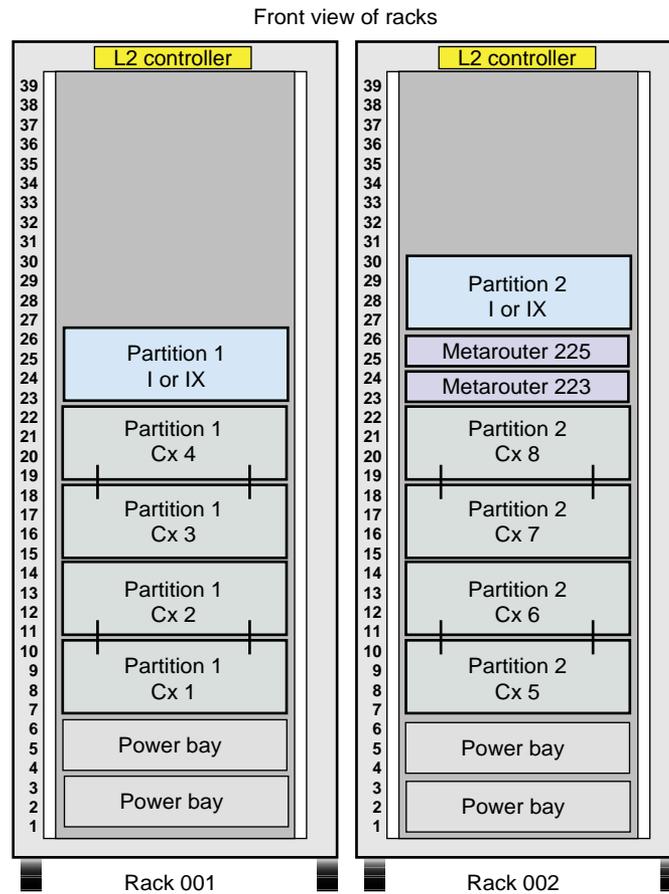


Figure 2-18 Brick Placement for Two 4-brick Partitions

One 4-brick Partition and Two 2-brick Partitions

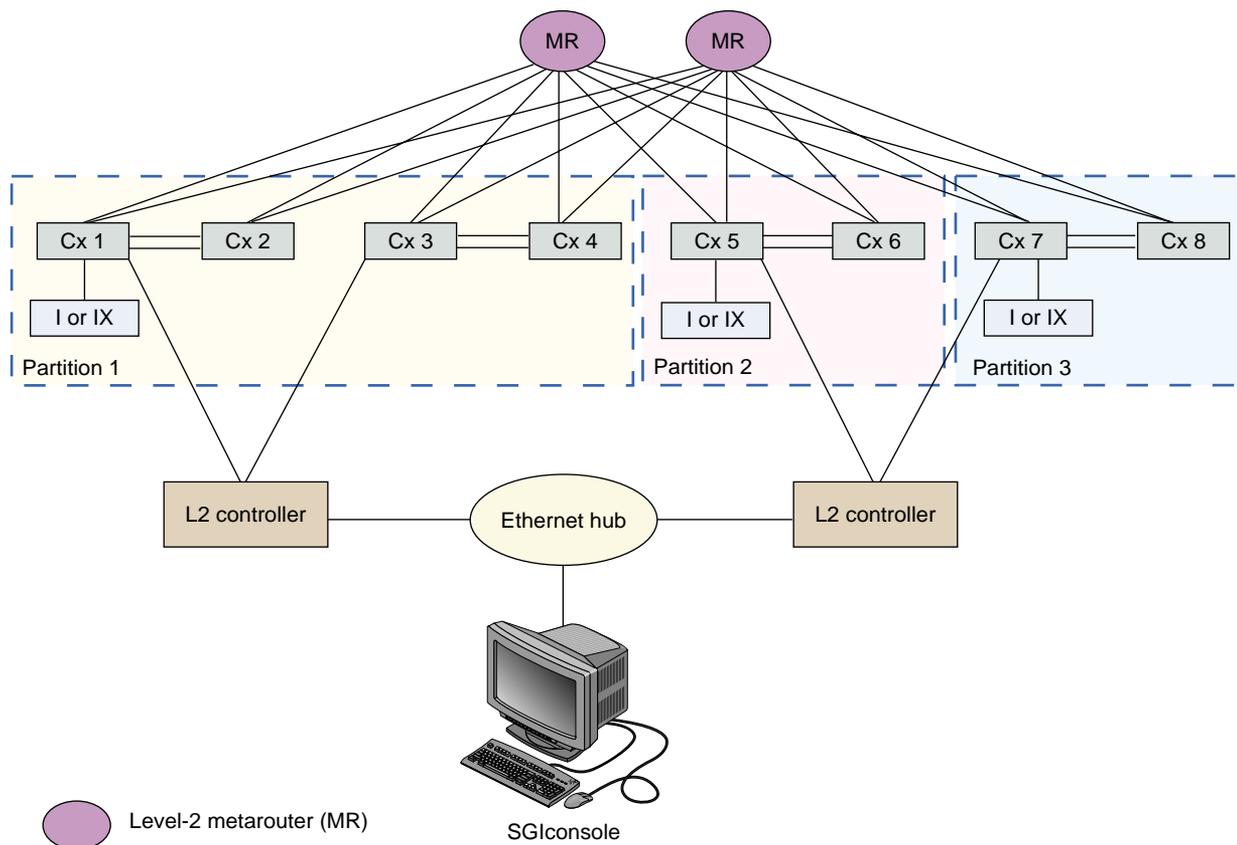


Figure 2-19 Example of One 4-brick Partition and Two 2-brick Partitions

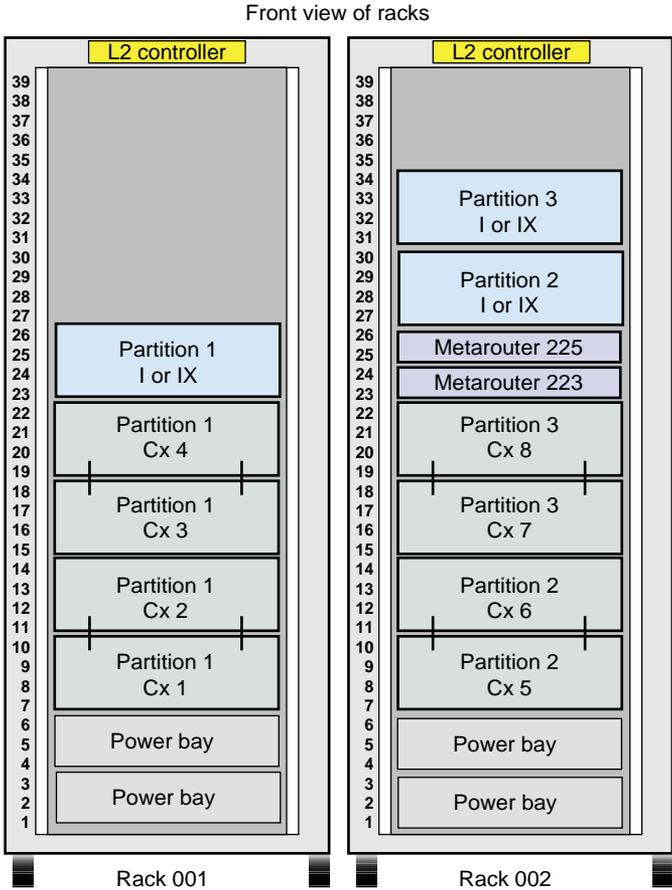
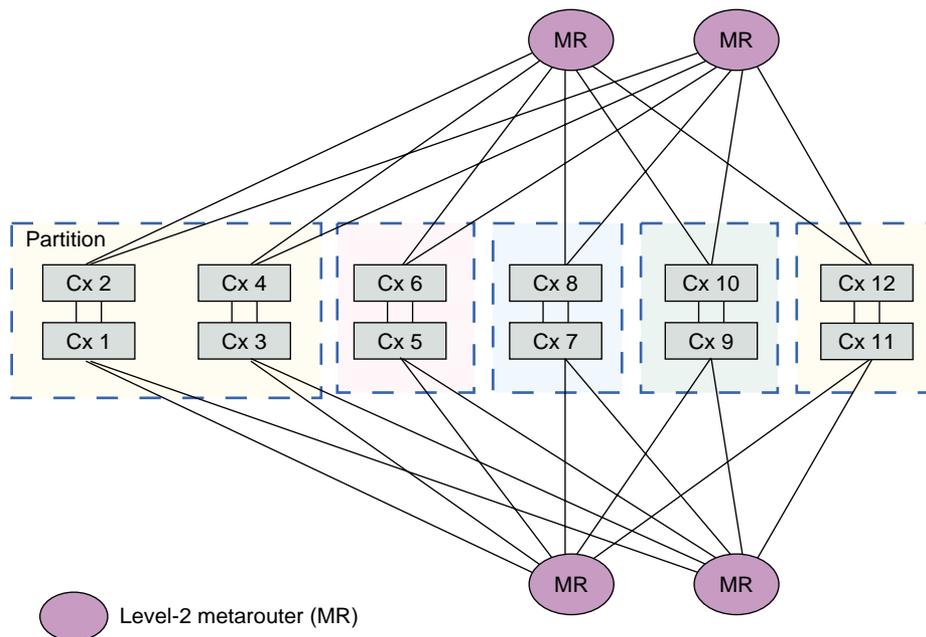


Figure 2-20 Brick Placement for One 4-brick Partition and Two 2-brick Partitions

One 4-brick Partition and Four 2-brick Partitions



Note: This illustration does not include the partition numbering, IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-21 Example of One 4-brick Partition and Four 2-brick Partitions

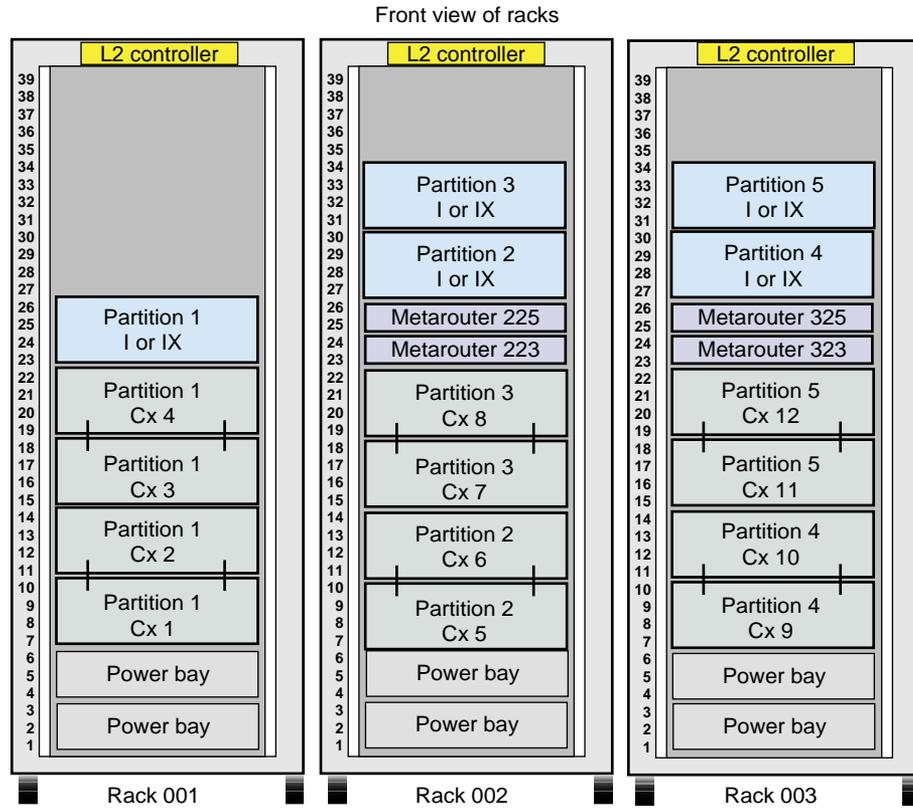
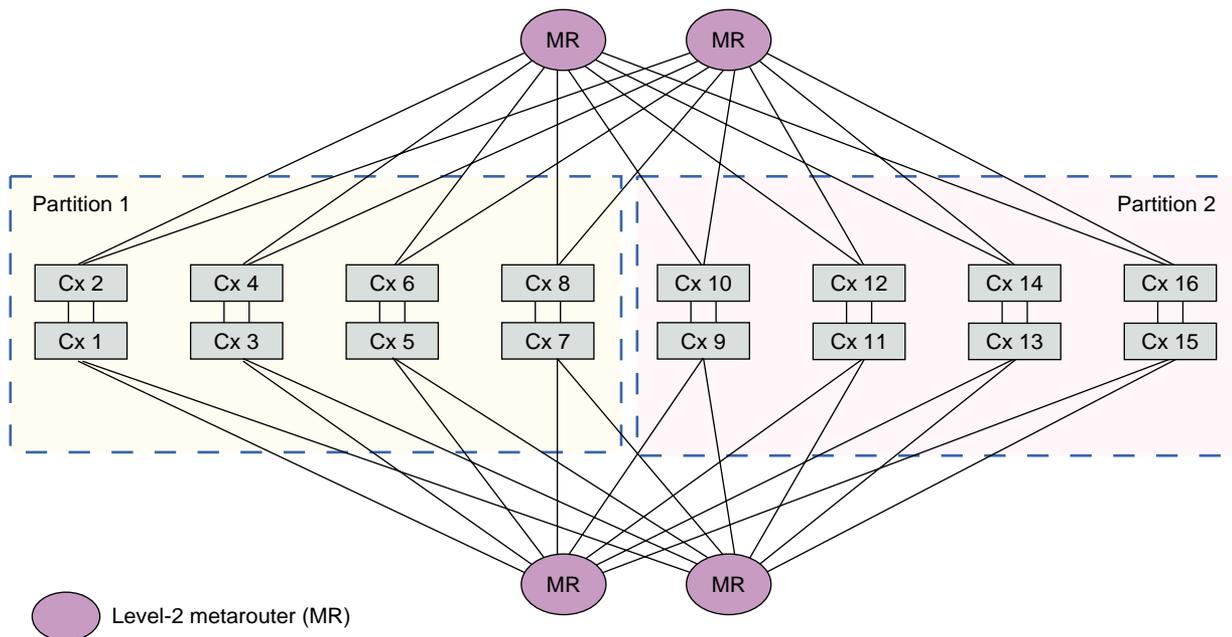


Figure 2-22 Brick Placement for One 4-brick Partition and Four 2-brick Partitions

Two 8-brick Partitions



Note: This illustration does not include the IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-23 Example of Two 8-brick Partitions

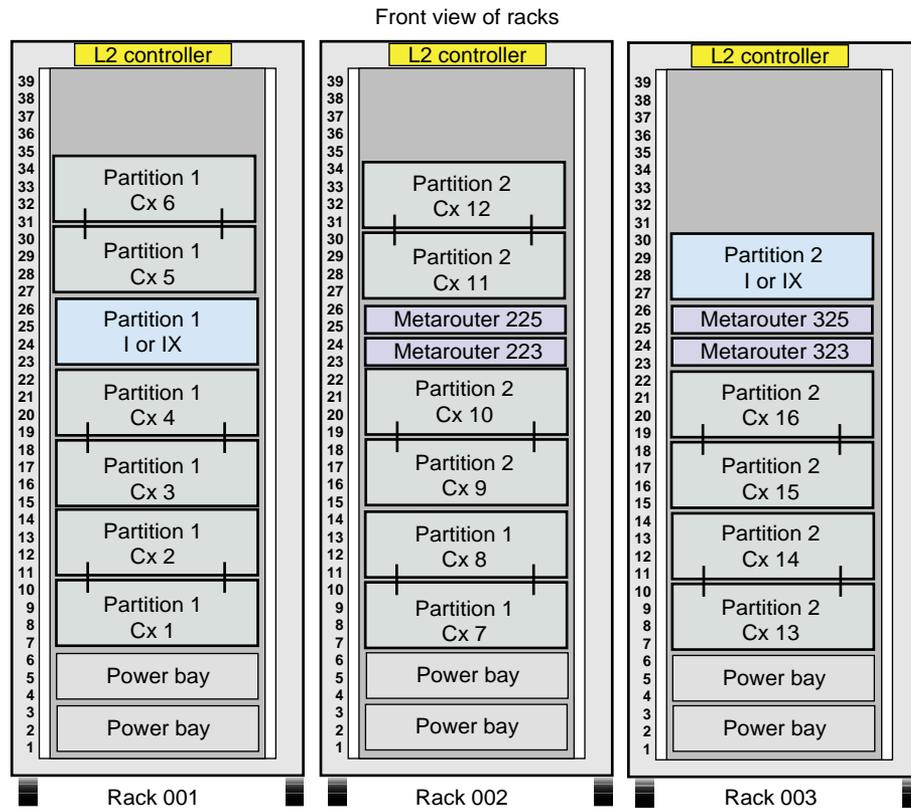
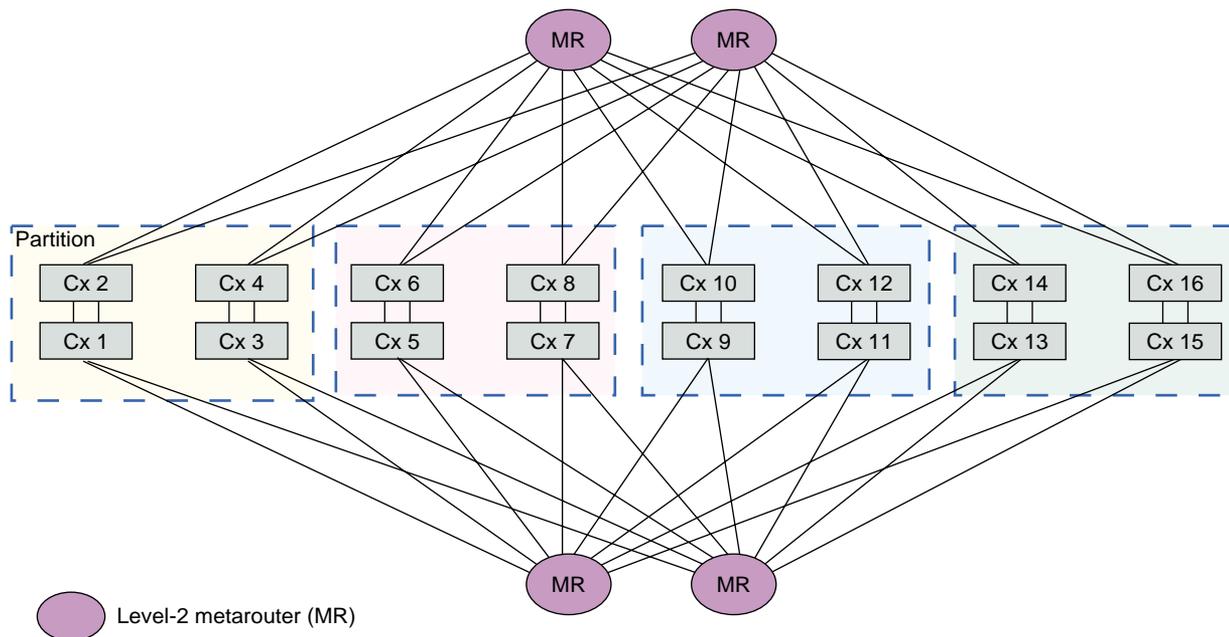


Figure 2-24 Brick Placement for Two 8-brick Partitions

Four 4-brick Partitions



Note: This illustration does not include the partition numbering, IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-25 Example of Four 4-brick Partitions

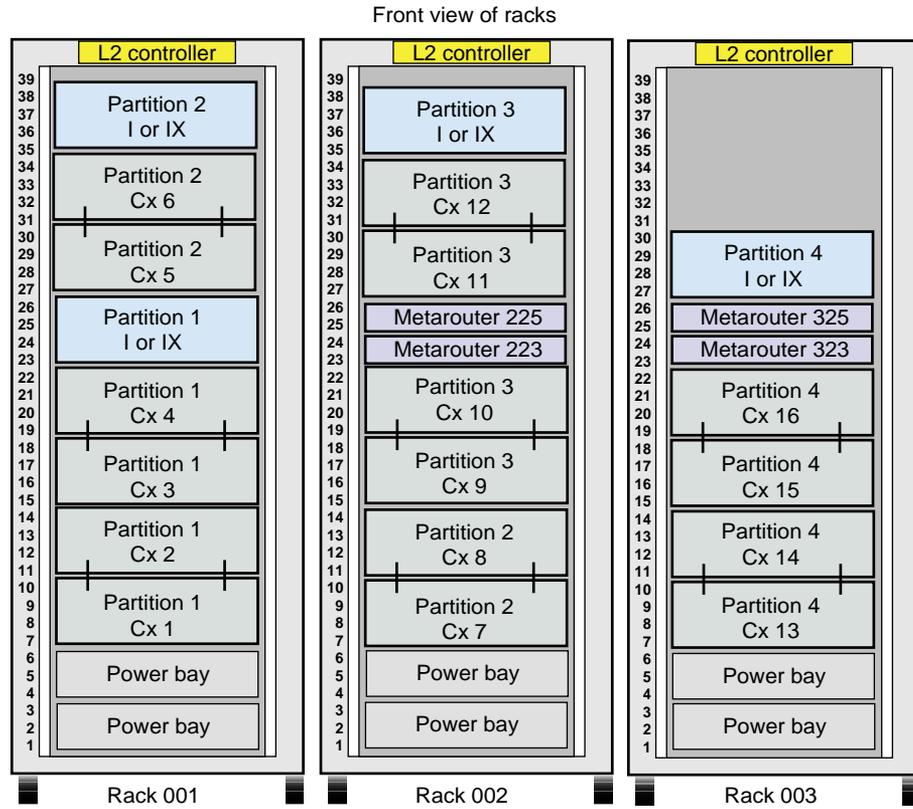
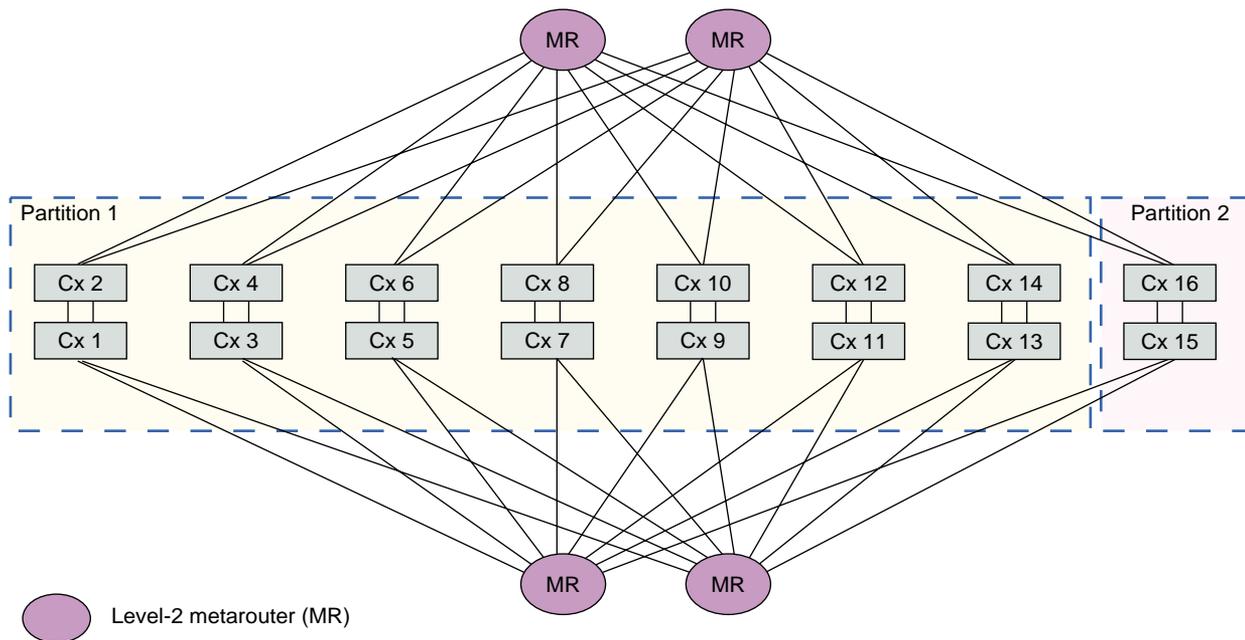


Figure 2-26 Brick Placement for Four 4-brick Partitions

One 14-brick Partition and One 2-brick Partition



Note: This illustration does not include the IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-27 Example of One 14-brick Partition and One 2-brick Partition

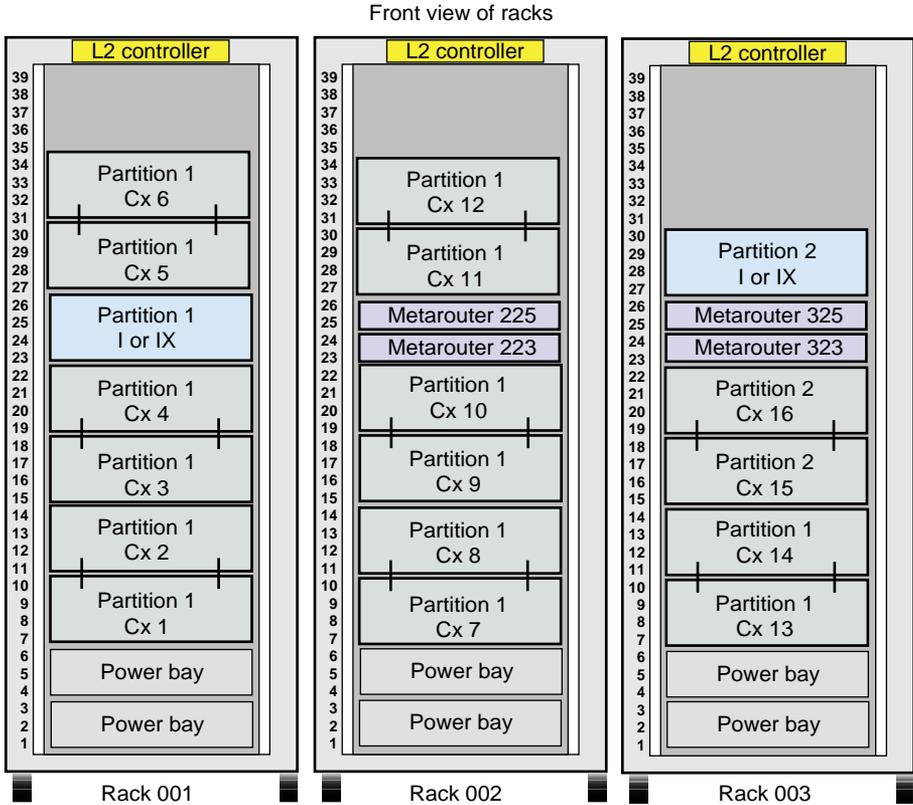
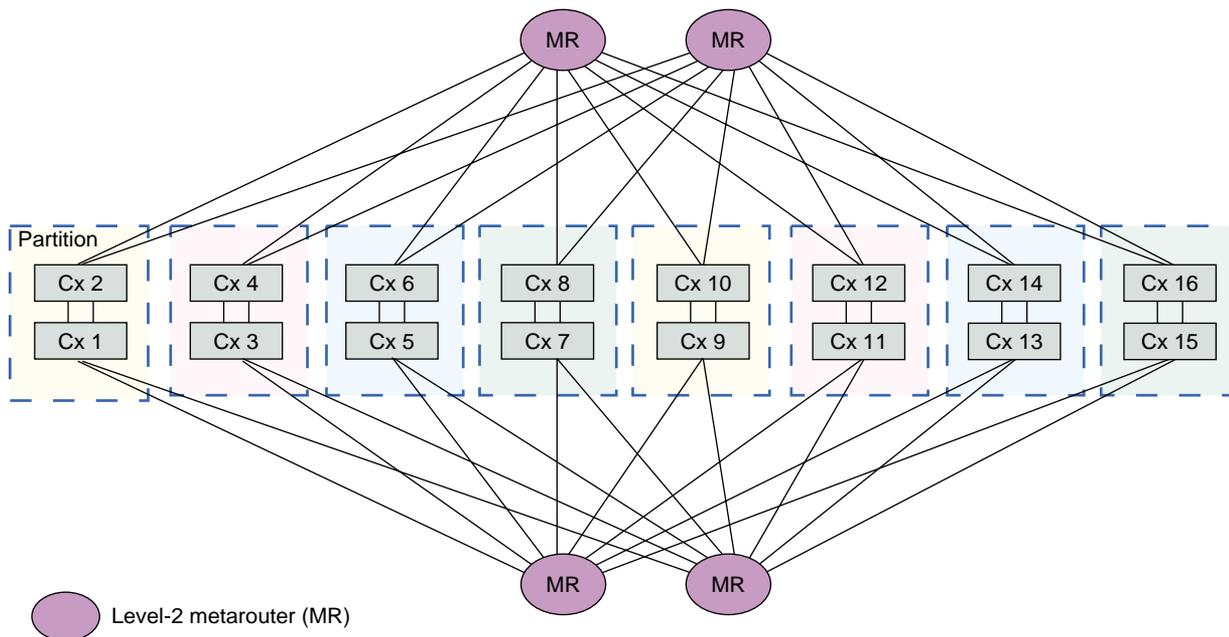


Figure 2-28 Brick Placement for One 14-brick Partition and One 2-brick Partition

Eight 2-brick Partitions



Note: This illustration does not include the partition numbering, IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-29 Example of Eight 2-brick Partitions

Front view of racks

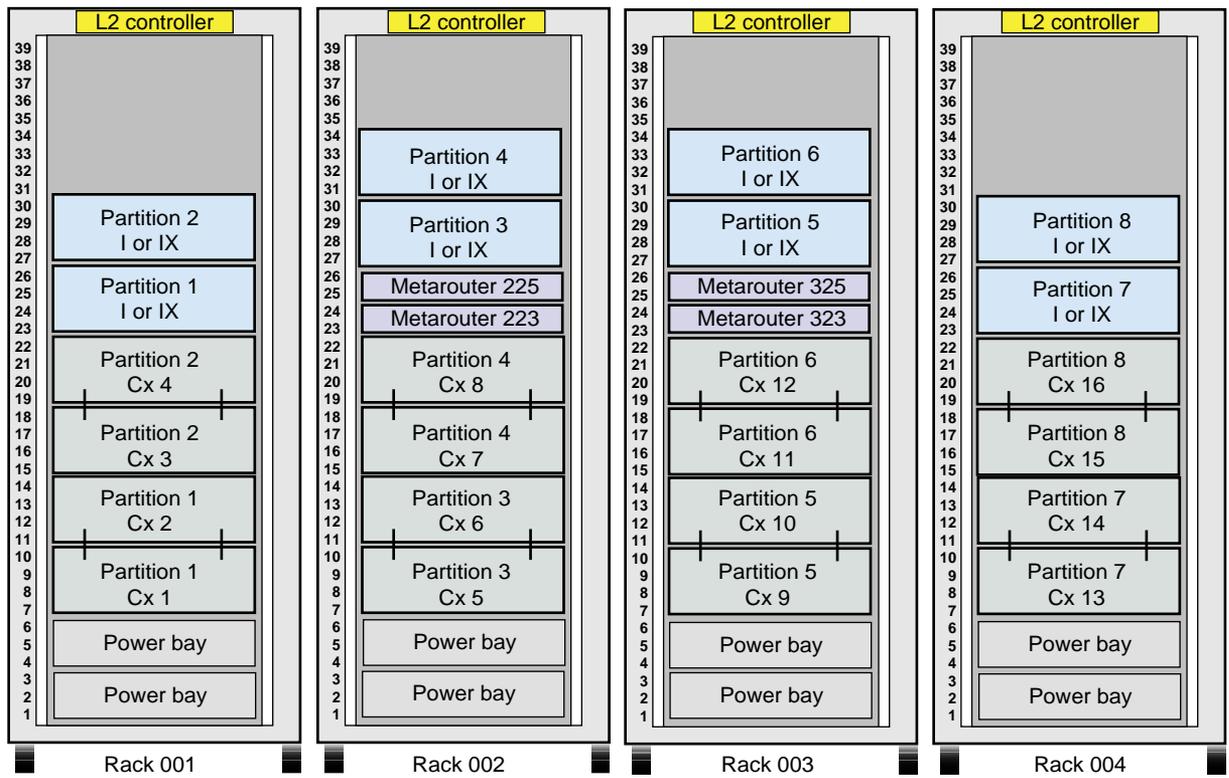
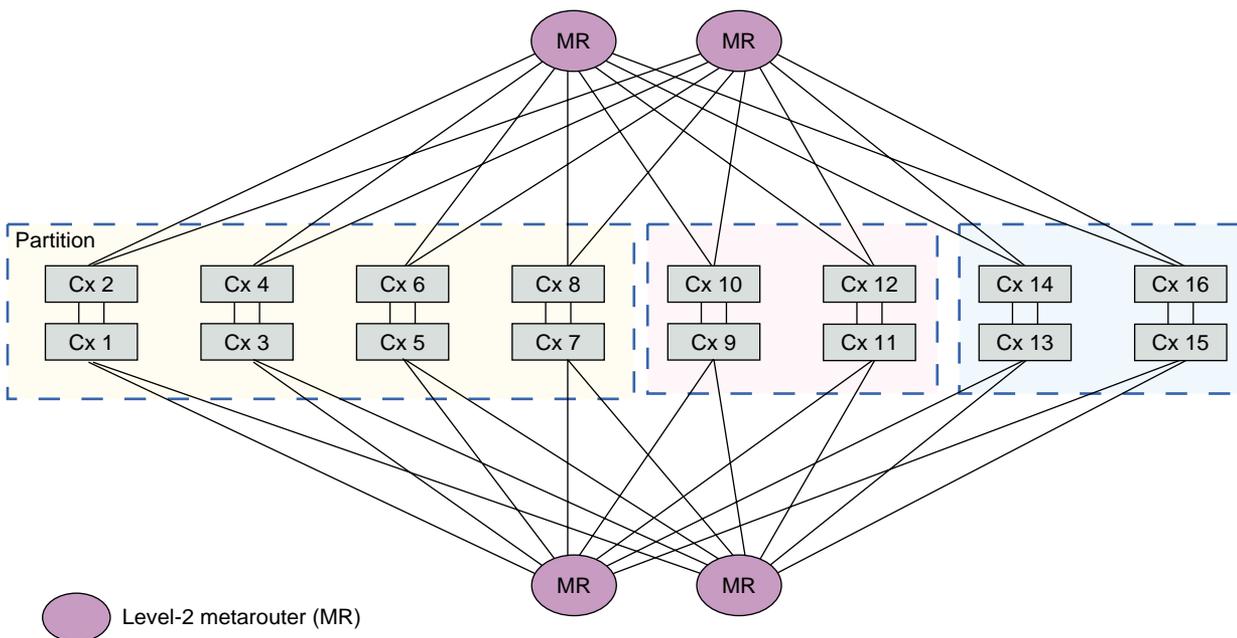


Figure 2-30 Brick Placement for Eight 2-brick Partitions

One 8-brick Partition and Two 4-brick Partitions



Note: This illustration does not include the partition numbering, IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-31 Example of One 8-brick Partition and Two 4-brick Partitions

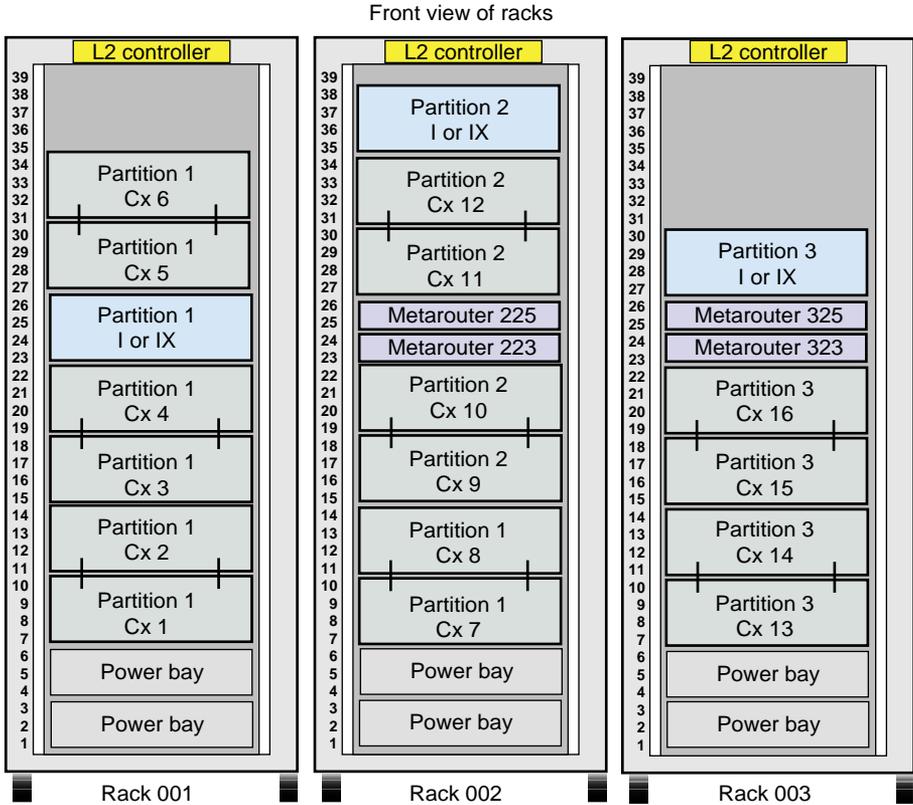
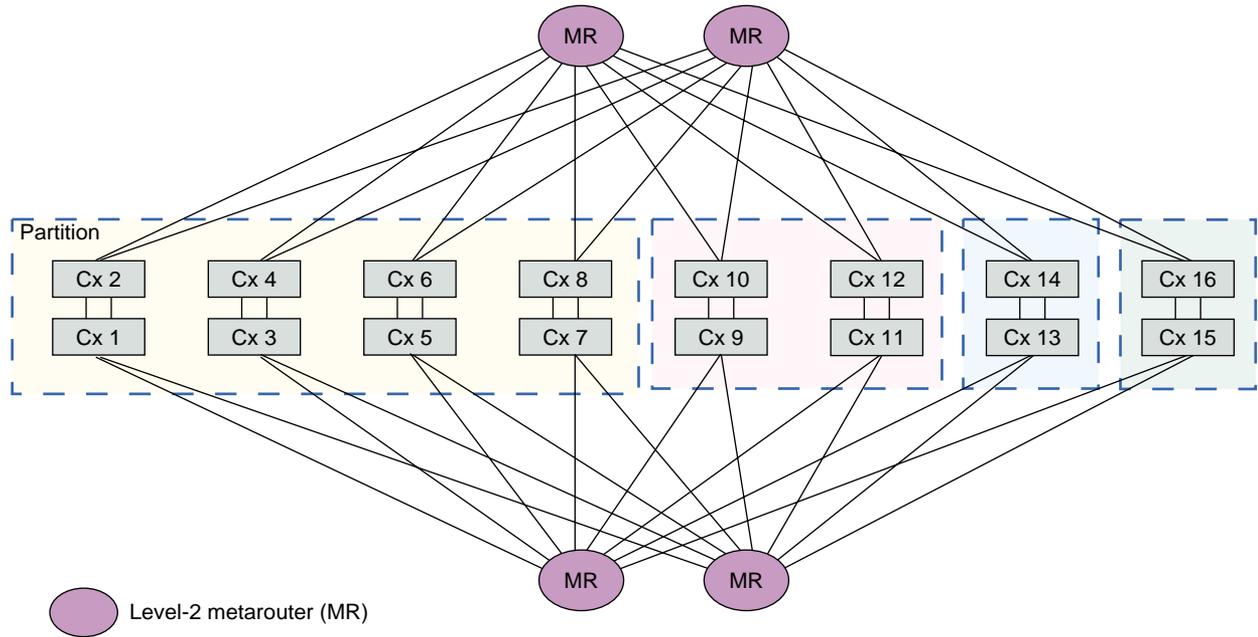


Figure 2-32 Brick Placement for One 8-brick Partition and Two 4-brick Partitions

One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions



Note: This illustration does not include the partition numbering, IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-33 Example of One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions

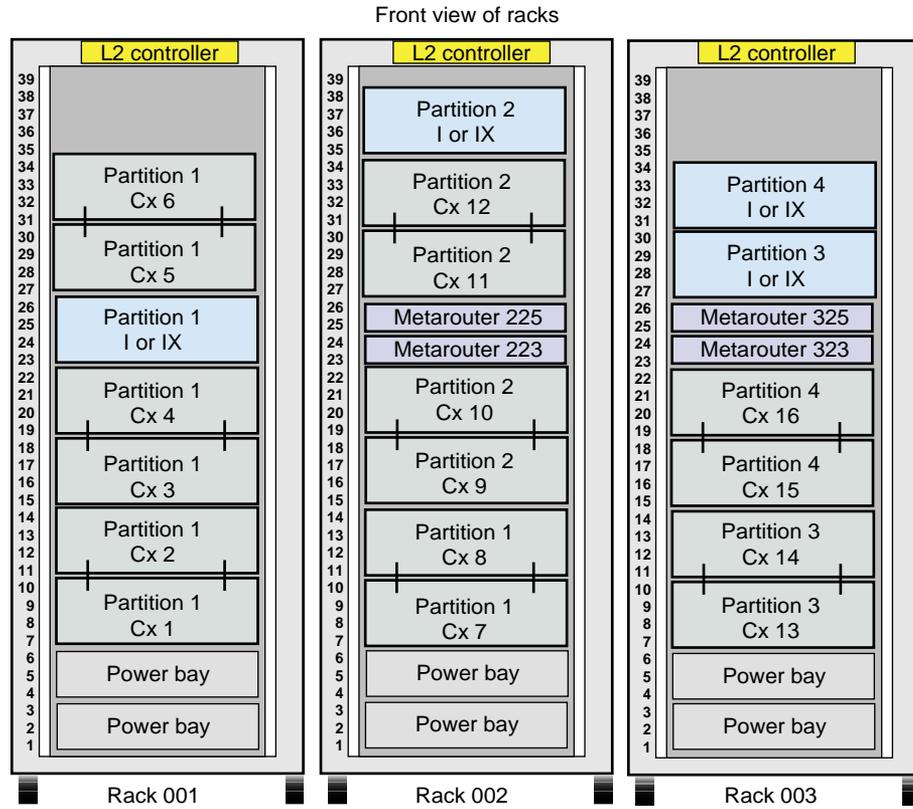
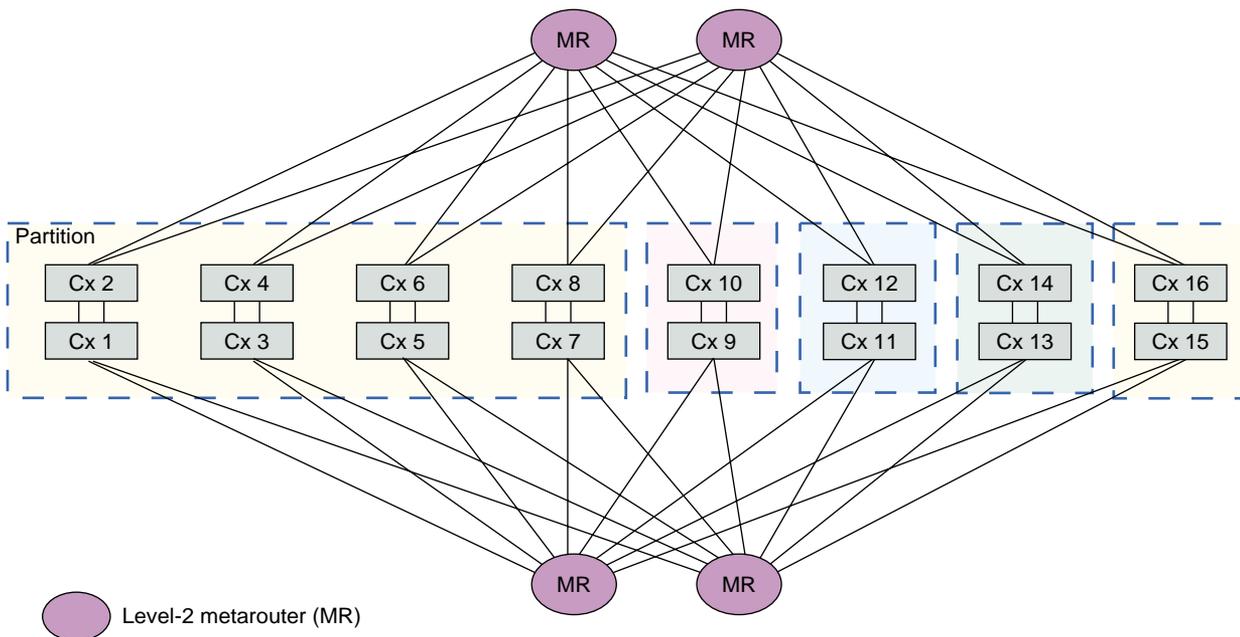


Figure 2-34 Brick Placement for One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions

One 8-brick Partition and Four 2-brick Partitions



Note: This illustration does not include the partition numbering, IX-bricks, L2 controller, Ethernet hub, and SGIconsole.

Figure 2-35 Example of One 8-brick Partition and Four 2-brick Partitions

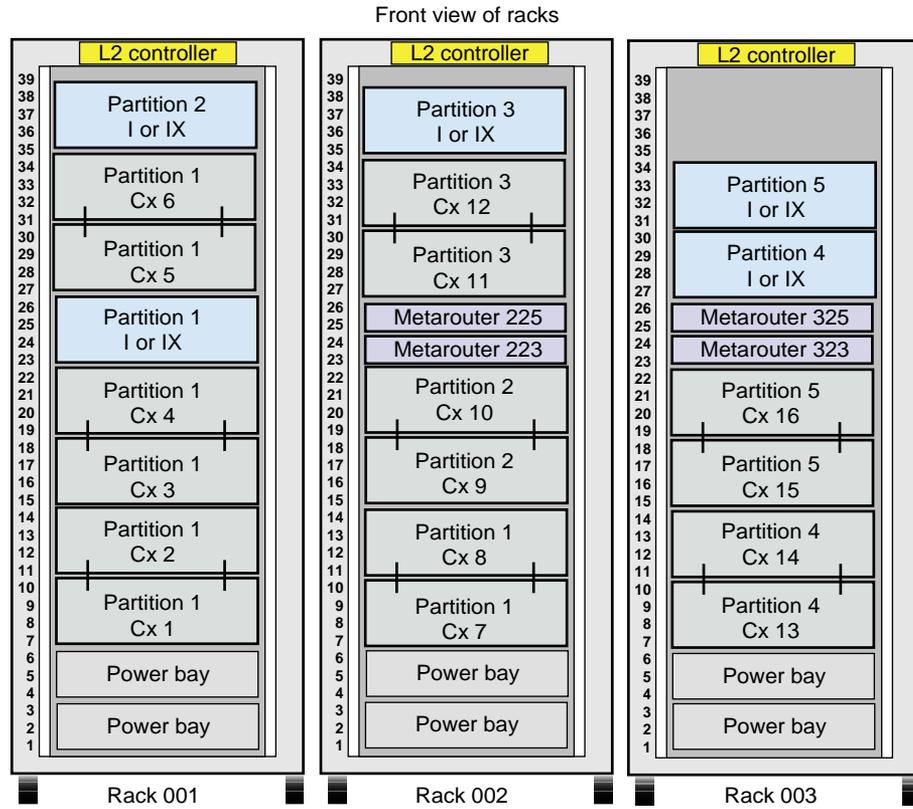


Figure 2-36 Brick Placement for One 8-brick Partition and Four 2-brick Partitions

Two 16-brick Partitions

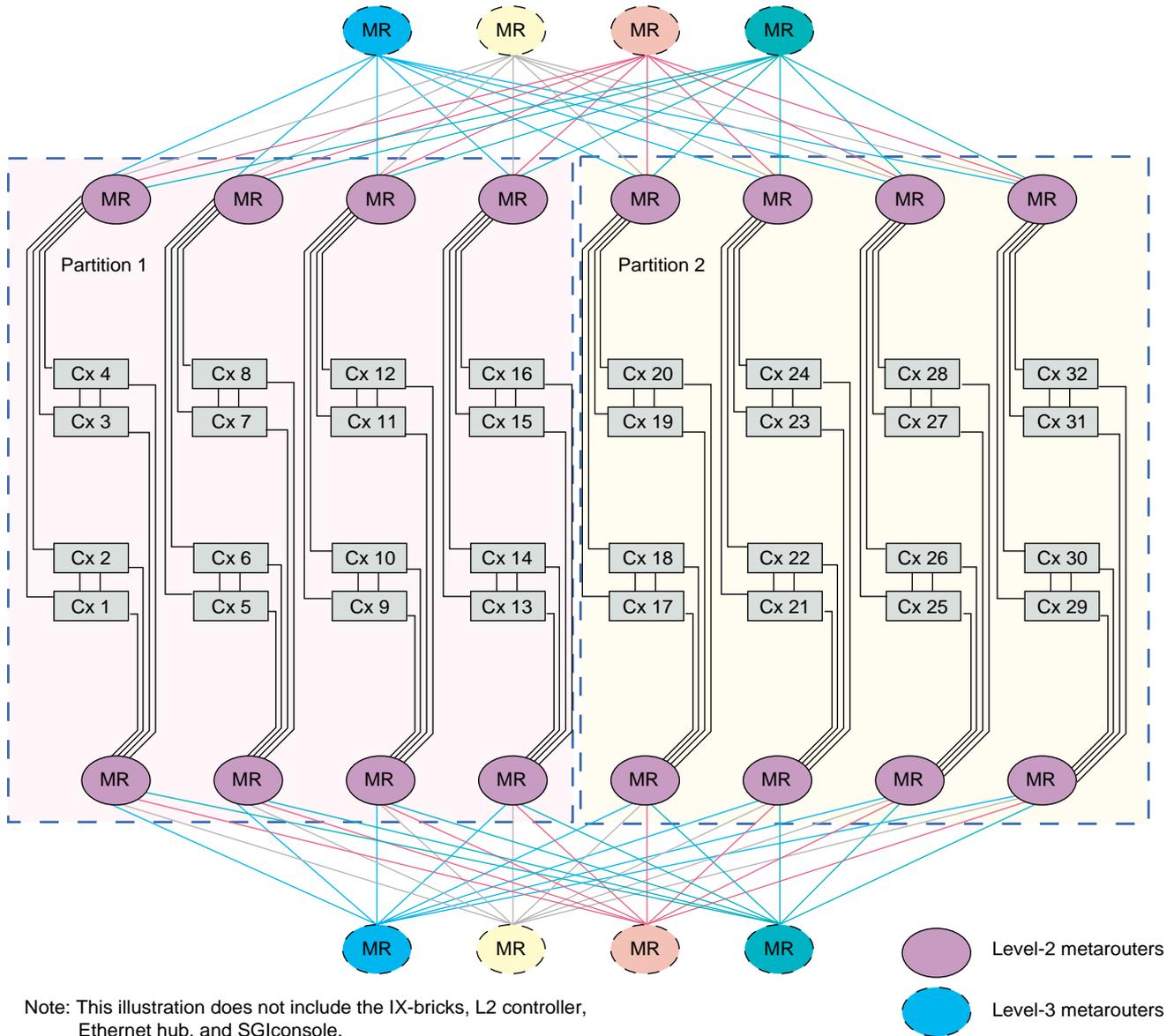


Figure 2-37 Example of Two 16-brick Partitions

Front view of racks

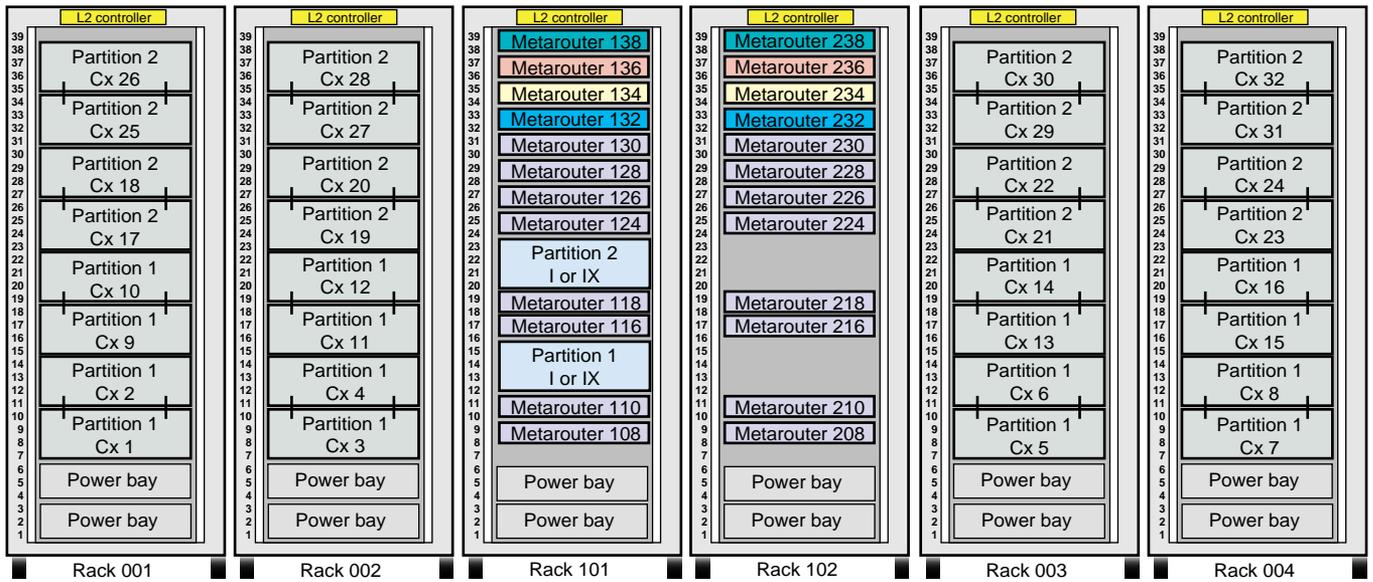


Figure 2-38 Brick Placement for Two 16-brick Partitions

Four 8-brick Partitions

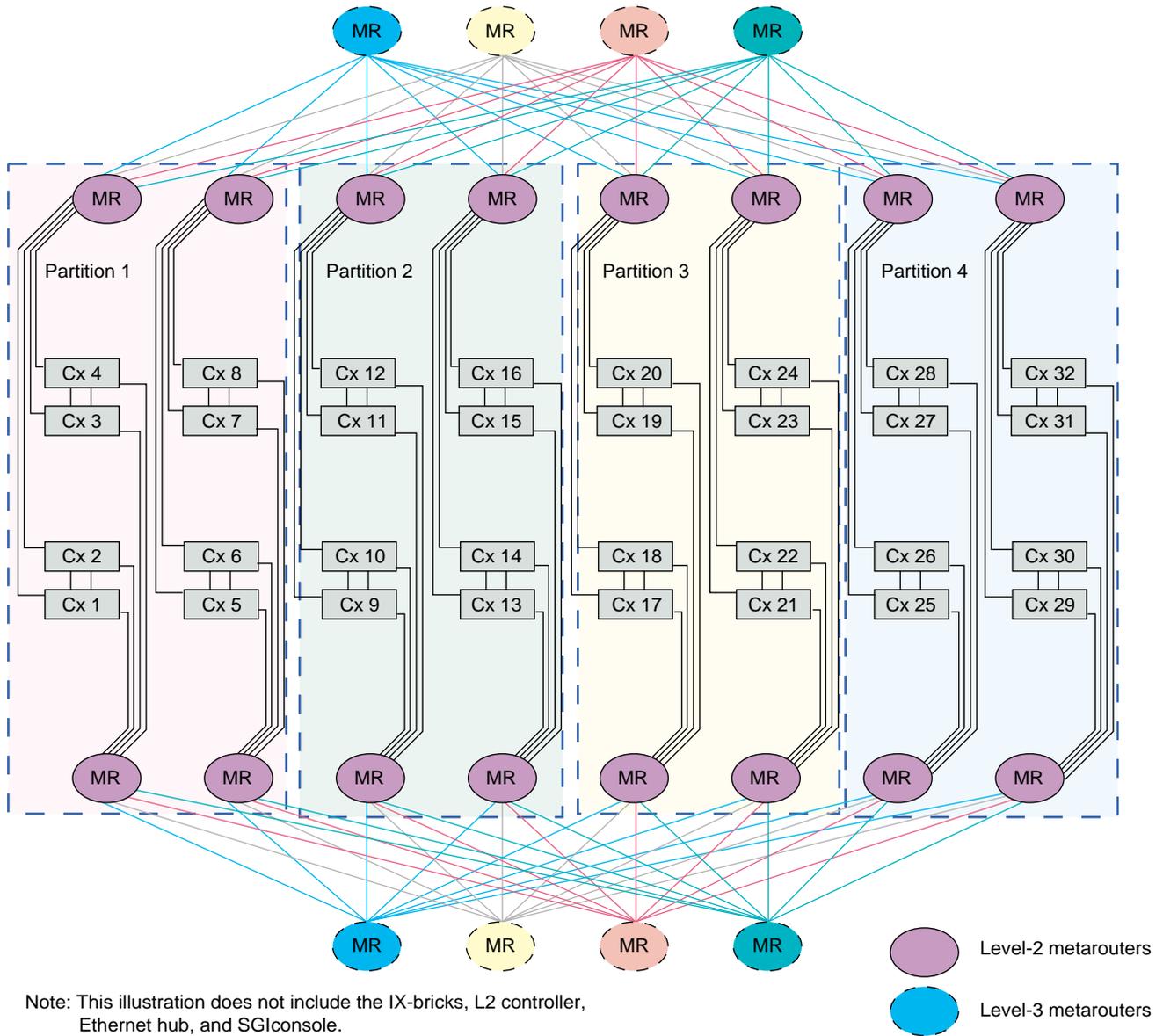


Figure 2-39 Example of Four 8-brick Partitions

Front view of racks

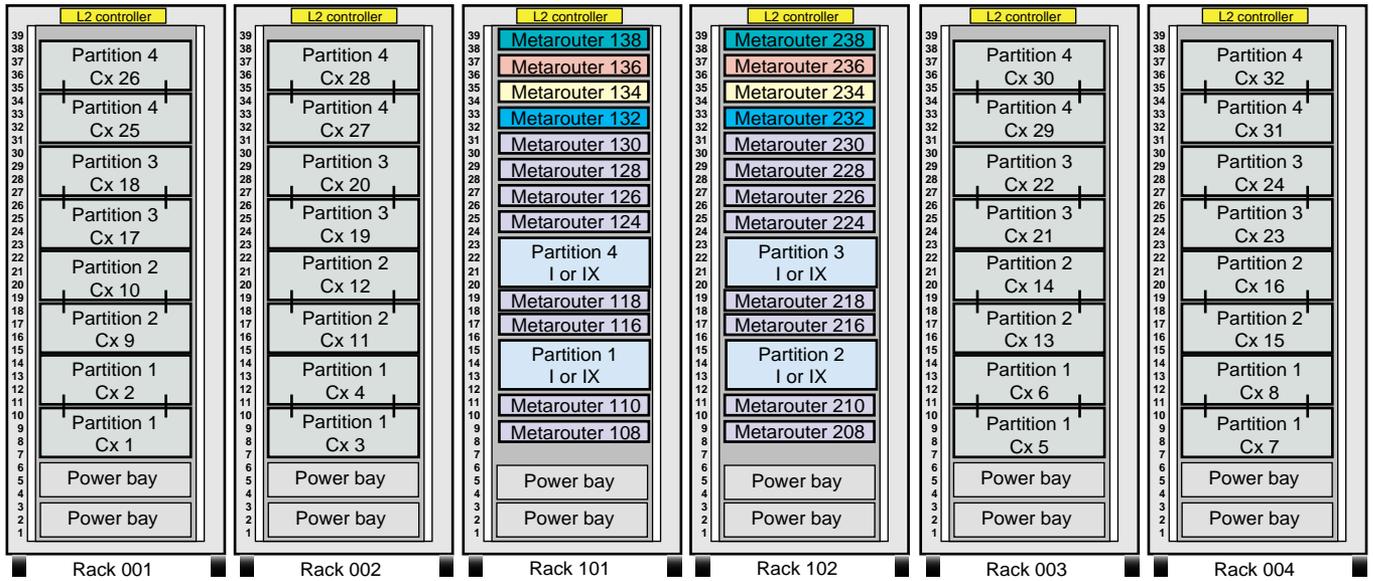


Figure 2-40 Brick Placement for Four 8-brick Partitions

Eight 4-brick Partitions

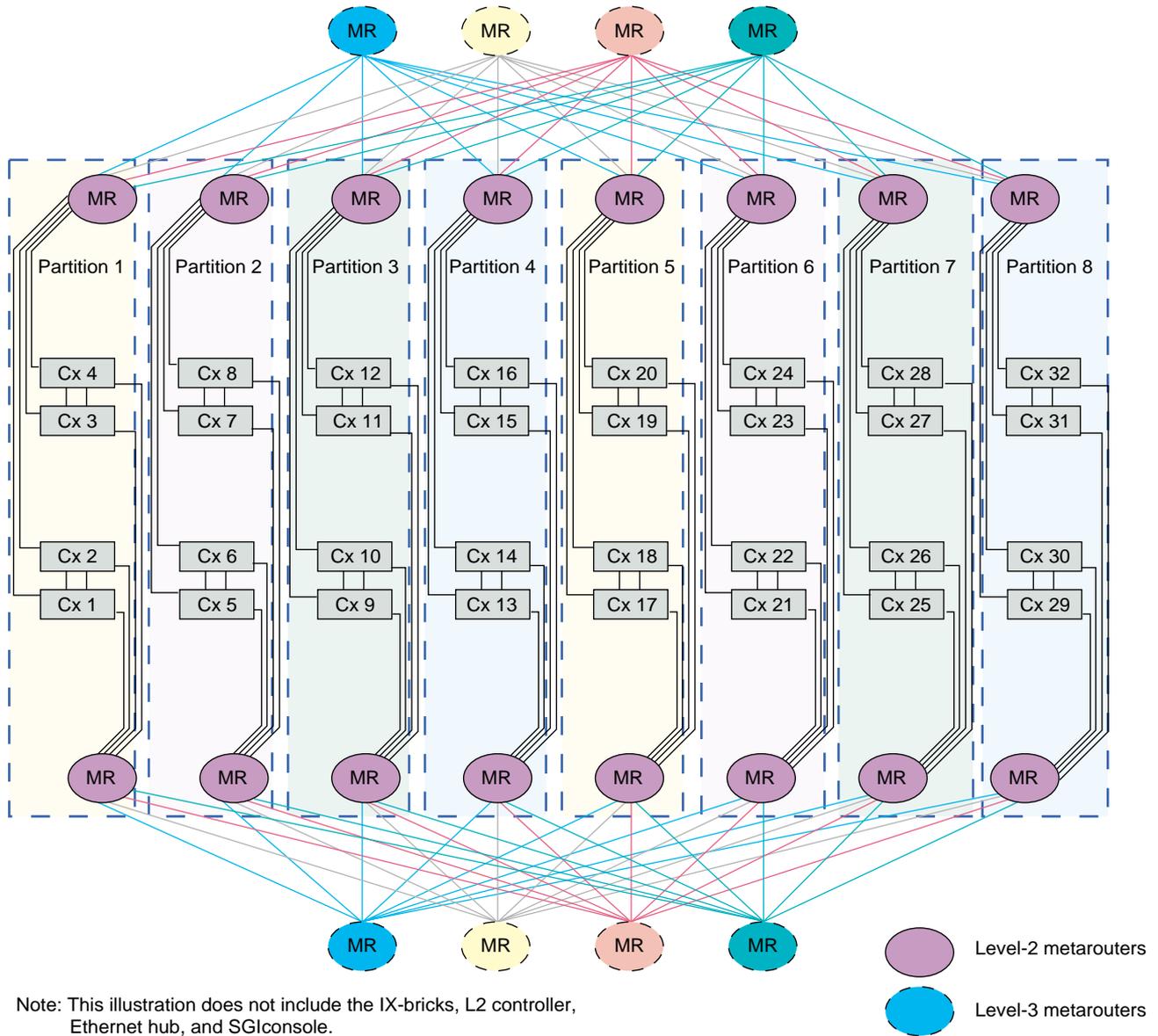


Figure 2-41 Example of Eight 4-brick Partitions

Front view of racks

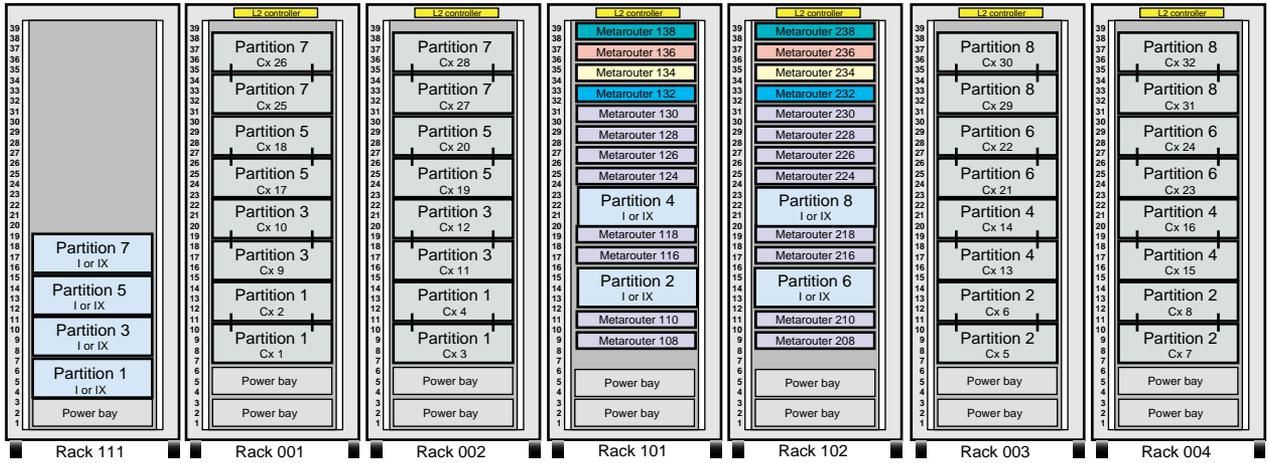


Figure 2-42 Brick Placement for Eight 4-brick Partitions

One 16-brick Partition and Two 8-brick Partitions

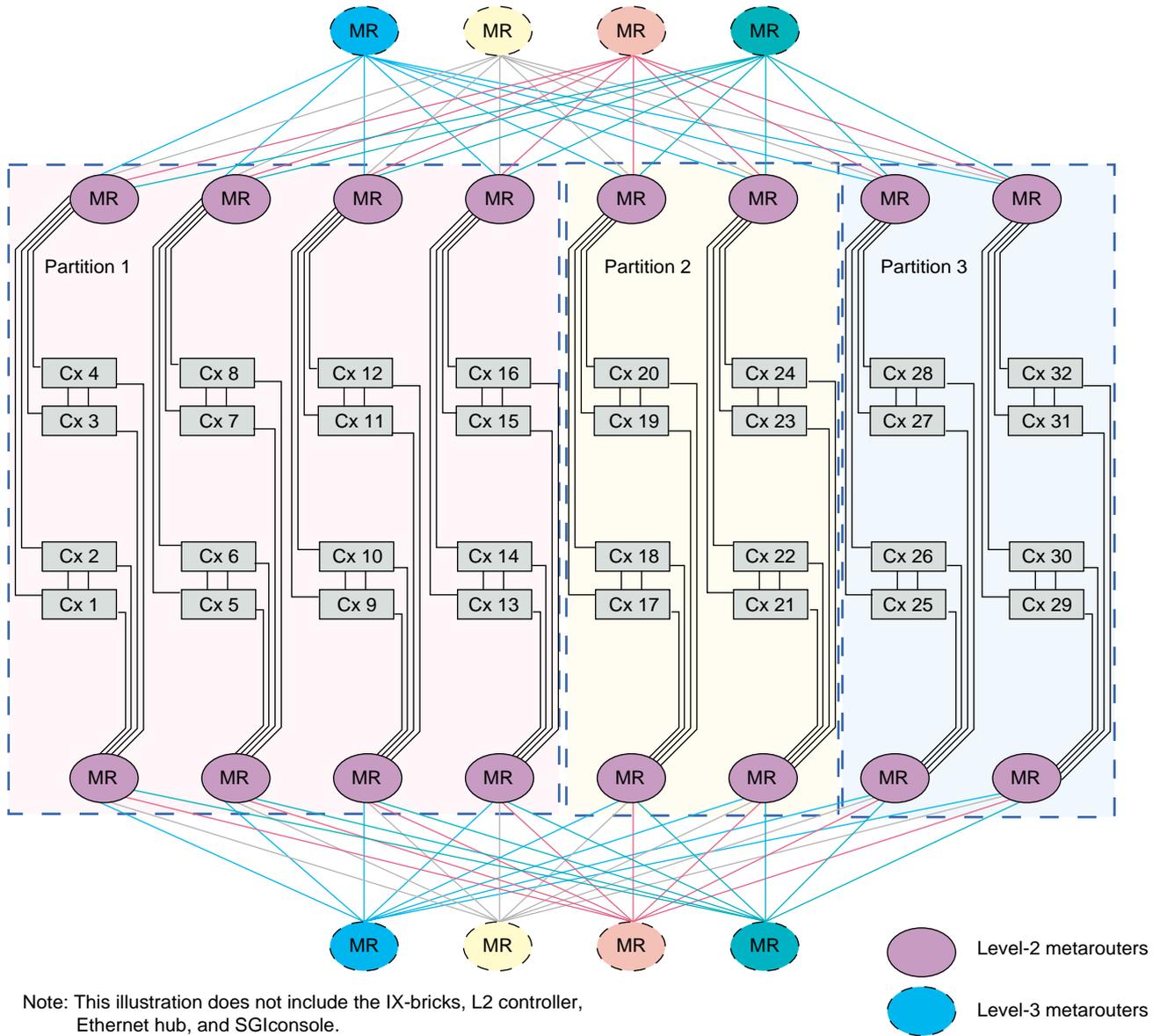


Figure 2-43 Example of One 16-brick Partition and Two 8-brick Partitions

One 16-brick Partition, One 8-brick Partition, and Two 4-brick Partitions

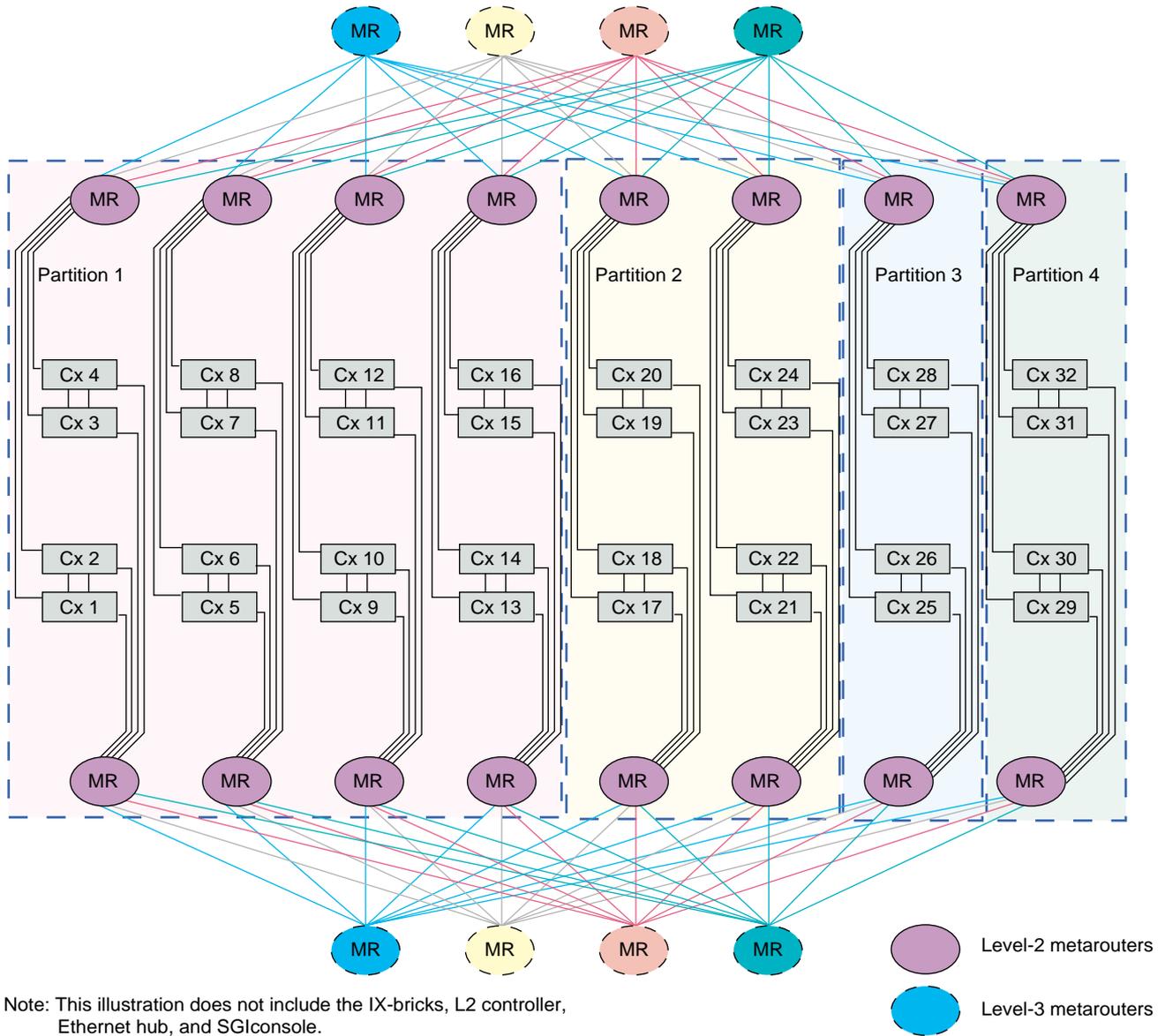


Figure 2-45 Example of One 16-brick Partition, One 8-brick Partition, and Two 4-brick Partitions

Front view of racks

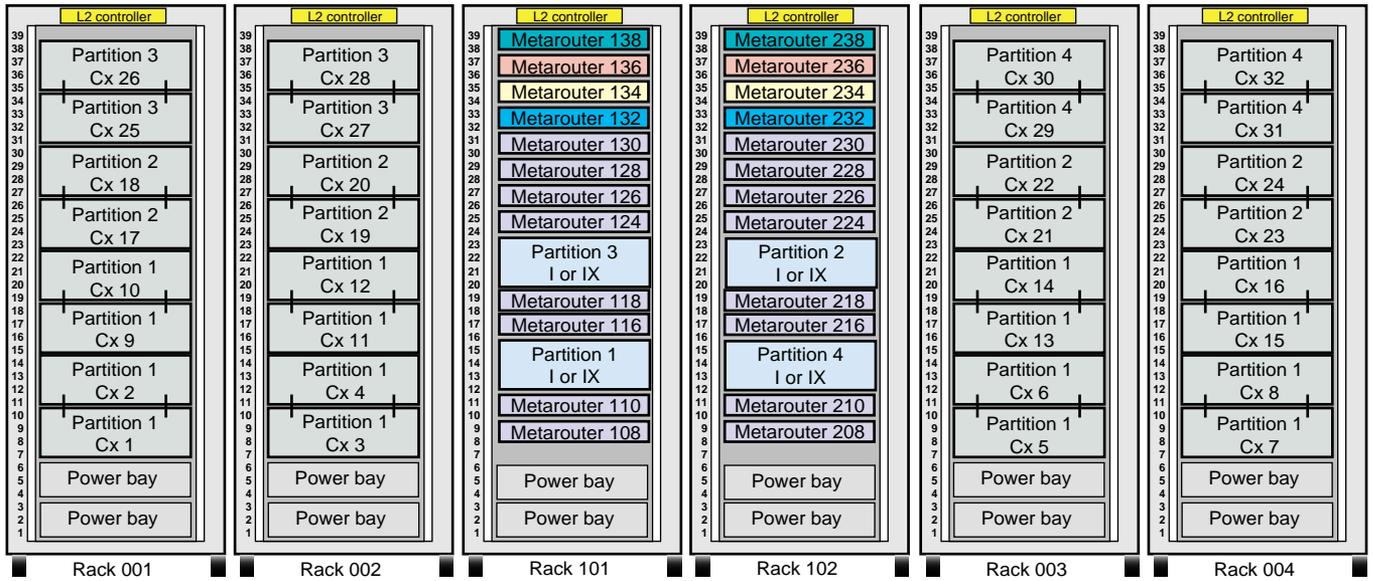


Figure 2-46 Brick Placement for One 16-brick Partition, One 8-brick Partition, and Two 4-brick Partitions

One 16-brick Partition, One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions

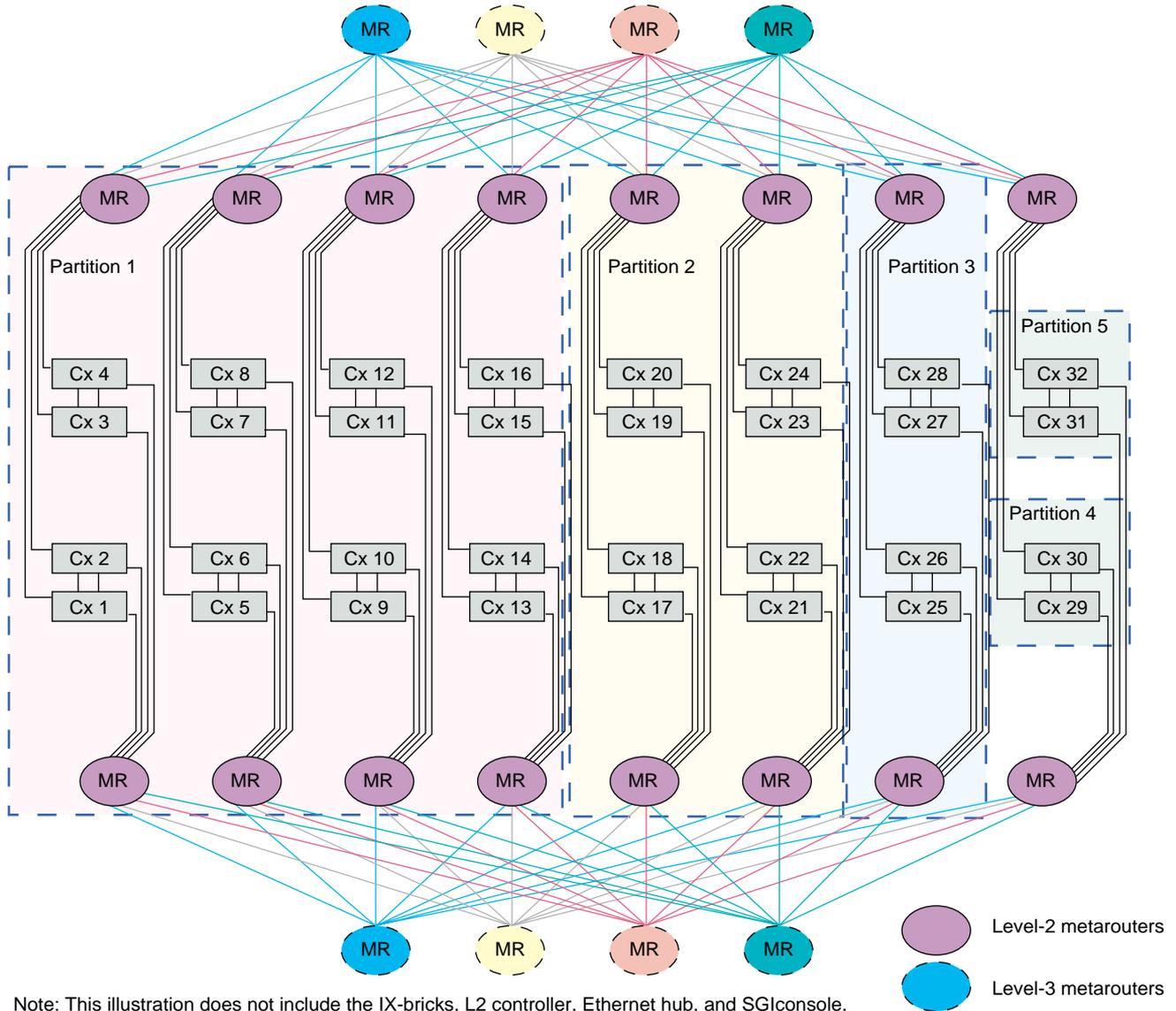


Figure 2-47 Example of One 16-brick Partition, One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions

Front view of racks

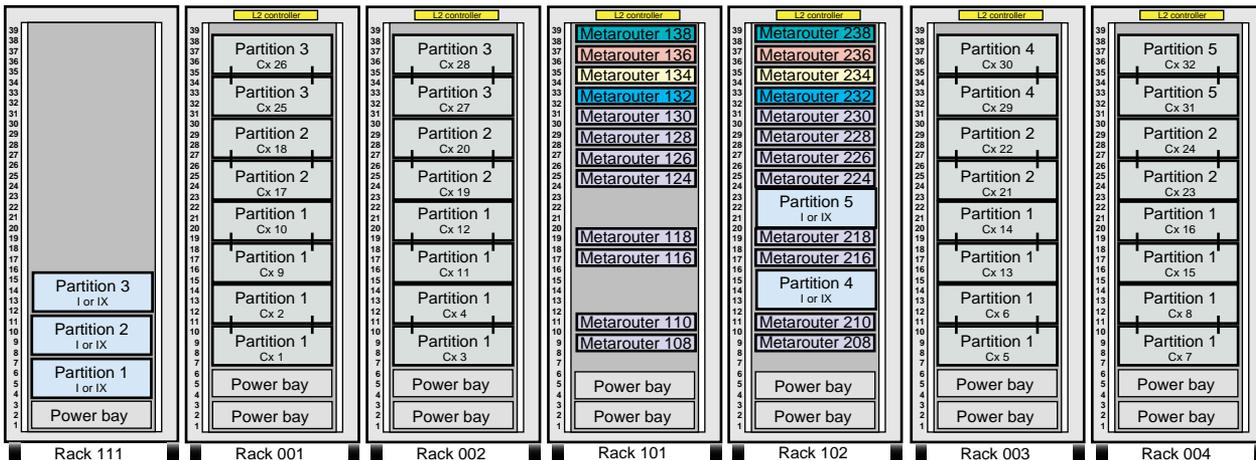


Figure 2-48 Brick Placement for One 16-brick Partition, One 8-brick Partition, One 4-brick Partition, and Two 2-brick Partitions

