



Cisco UCS B480 M5 Blade Server Installation and Service Note

First Published: 2017-10-28

Last Modified: 2020-09-08

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883



CONTENTS

Introduction vii

PREFACE

Preface ix

Audience ix

Conventions ix

Related Cisco UCS Documentation xi

Obtaining Documentation and Submitting a Service Request xi

CHAPTER 1

Overview 1

Cisco UCS B480 M5 Blade Server 1

External Features Overview 3

LEDs 3

Buttons 4

Local Console Connection 5

Drive Bays 5

Graphics Processing Unit 5

CHAPTER 2

Installing a Blade Server 7

Installing a Blade Server 7

Server Configuration 9

Powering Off a Blade Server Using the Power Button 9

Removing a Blade Server 10

Server Troubleshooting 10

CHAPTER 3

Servicing a Blade Server 11

Replacing a Drive 11

Removing a Blade Server Drive	12
Installing a Blade Server Drive	12
4K Sector Format SAS/SATA Drives Considerations	13
Setting Up UEFI Mode Booting in the UCS Manager Boot Policy	13
Removing a Blade Server Cover	14
Internal Components	14
Diagnostics Button and LEDs	16
Installing the Front Mezzanine Storage Module	16
Replacing the SuperCap Module	17
Removing the SuperCap Module	17
Installing the SuperCap Module	19
Replacing CPUs and Heatsinks	20
Special Information For Upgrades to Second Generation Intel Xeon Scalable Processors	20
CPU Configuration Rules	21
Tools Required for CPU Replacement	22
Replacing a CPU and Heatsink	23
Additional CPU-Related Parts to Order with CPU RMA	29
Additional CPU-Related Parts to Order with RMA Replacement Blade Server	30
Moving an M5 Generation CPU	30
Replacing Memory DIMMs	35
Memory Population Guidelines	35
Installing a DIMM or DIMM Blank	37
Memory Performance	38
Memory Mirroring and RAS	38
Replacing Intel Optane DC Persistent Memory Modules	39
Intel Optane DC Persistent Memory Module Population Rules and Performance Guidelines	39
Installing Intel Optane DC Persistent Memory Modules	41
Server BIOS Setup Utility Menu for DCPMM	43
Installing a Virtual Interface Card in the mLOM Slot	44
Installing a Rear Mezzanine Module in Addition to the mLOM VIC	45
NVIDIA P6 Graphics Processing Unit	46
Installing an NVIDIA GPU Card in the Front of the Server	47
Installing an NVIDIA GPU Card in the Rear of the Server	51
Enabling the Trusted Platform Module	52

Removing the Trusted Platform Module (TPM)	54
Mini Storage Module	55
Replacing the Mini-Storage Module Carrier	56
Replacing an M.2 SSD	57
Embedded SATA RAID Controller	58
Embedded SATA RAID Requirements	58
Embedded SATA RAID Controller Considerations	59
Embedded SATA RAID: Two SATA Controllers	59
Enabling SATA Mode	59
Accessing the LSI Software RAID Configuration Utility	60
Installing LSI MegaSR Drivers For Windows and Linux	60
For More RAID Utility Information	67
Replacing a Boot-Optimized M.2 RAID Controller Module	67
Cisco Boot-Optimized M.2 RAID Controller Considerations	67
Replacing a Cisco Boot-Optimized M.2 RAID Controller	69
Recycling the PCB Assembly (PCBA)	71

CHAPTER 4 **Technical Specifications** **73**

Physical Specifications for the UCS B480 Blade Server	73
---	----

CHAPTER 5 **NVIDIA Licensing Information** **75**

NVIDIA GRID License Server Overview	75
Registering Your Product Activation Keys with NVIDIA	76
Downloading the GRID Software Suite	77
Installing NVIDIA GRID License Server Software	77
Platform Requirements for NVIDIA GRID License Server	77
Installing on Windows	78
Installing on Linux	78
Installing GRID Licenses From the NVIDIA Licensing Portal to the License Server	79
Accessing the GRID License Server Management Interface	79
Reading Your License Server's MAC Address	79
Installing Licenses From the Licensing Portal	80
Viewing Available Licenses	80
Viewing Current License Usage	80

Managing GRID Licenses	81
Acquiring a GRID License on Windows	81
Acquiring a GRID License on Linux	82
Installing Drivers to Support the NVIDIA GPU Cards	82
1. Updating the Server BIOS Firmware	83
2. Activating the Server BIOS Firmware	83
3. Updating the NVIDIA Drivers	84

Introduction

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The following information is for FCC compliance of Class A devices: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to part 15 of the FCC rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio-frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference, in which case users will be required to correct the interference at their own expense.

The following information is for FCC compliance of Class B devices: This equipment has been tested and found to comply with the limits for a Class B digital device, pursuant to part 15 of the FCC rules. These limits are designed to provide reasonable protection against harmful interference in a residential installation. This equipment generates, uses and can radiate radio frequency energy and, if not installed and used in accordance with the instructions, may cause harmful interference to radio communications. However, there is no guarantee that interference will not occur in a particular installation. If the equipment causes interference to radio or television reception, which can be determined by turning the equipment off and on, users are encouraged to try to correct the interference by using one or more of the following measures:

- Reorient or relocate the receiving antenna.
- Increase the separation between the equipment and receiver.
- Connect the equipment into an outlet on a circuit different from that to which the receiver is connected.
- Consult the dealer or an experienced radio/TV technician for help.

Modifications to this product not authorized by Cisco could void the FCC approval and negate your authority to operate the product.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE

THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

All printed copies and duplicate soft copies of this document are considered uncontrolled. See the current online version for the latest version.

Cisco has more than 200 offices worldwide. Addresses and phone numbers are listed on the Cisco website at www.cisco.com/go/offices.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/c/en/us/about/legal/trademarks.html>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1721R)



Preface

- [Audience, on page ix](#)
- [Conventions, on page ix](#)
- [Related Cisco UCS Documentation, on page xi](#)
- [Obtaining Documentation and Submitting a Service Request, on page xi](#)

Audience

To use this installation guide, you must be familiar with electronic circuitry and wiring practices and preferably be an electronic or electromechanical technician who has experience with electronic and electromechanical equipment.

Only trained and qualified service personnel (as defined in IEC 60950-1 and AS/NZS60950) should install, replace, or service the equipment. Install the system in accordance with the U.S. National Electric Code if you are in the United States.

Conventions

Text Type	Indication
GUI elements	GUI elements such as tab titles, area names, and field labels appear in this font . Main titles such as window, dialog box, and wizard titles appear in this font .
Document titles	Document titles appear in <i>this font</i> .
TUI elements	In a Text-based User Interface, text the system displays appears in <i>this font</i> .
System output	Terminal sessions and information that the system displays appear in <i>this font</i> .
CLI commands	CLI command keywords appear in this font . Variables in a CLI command appear in <i>this font</i> .
[]	Elements in square brackets are optional.

Text Type	Indication
{x y z}	Required alternative keywords are grouped in braces and separated by vertical bars.
[x y z]	Optional alternative keywords are grouped in brackets and separated by vertical bars.
string	A nonquoted set of characters. Do not use quotation marks around the string or the string will include the quotation marks.
<>	Nonprinting characters such as passwords are in angle brackets.
[]	Default responses to system prompts are in square brackets.
!, #	An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line.



Note Means *reader take note*. Notes contain helpful suggestions or references to material not covered in the document.



Tip Means *the following information will help you solve a problem*. The tips information might not be troubleshooting or even an action, but could be useful information, similar to a Timesaver.



Timesaver Means *the described action saves time*. You can save time by performing the action described in the paragraph.



Caution Means *reader be careful*. In this situation, you might perform an action that could result in equipment damage or loss of data.



Warning IMPORTANT SAFETY INSTRUCTIONS

This warning symbol means danger. You are in a situation that could cause bodily injury. Before you work on any equipment, be aware of the hazards involved with electrical circuitry and be familiar with standard practices for preventing accidents. Use the statement number provided at the end of each warning to locate its translation in the translated safety warnings that accompanied this device.

SAVE THESE INSTRUCTIONS

Related Cisco UCS Documentation

Documentation Roadmaps

For a complete list of all B-Series documentation, see the *Cisco UCS B-Series Servers Documentation Roadmap* available at the following URL: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/overview/guide/UCS_roadmap.html

For a complete list of all C-Series documentation, see the *Cisco UCS C-Series Servers Documentation Roadmap* available at the following URL: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/overview/guide/ucs_rack_roadmap.html.

For information on supported firmware versions and supported UCS Manager versions for the rack servers that are integrated with the UCS Manager for management, refer to [Release Bundle Contents for Cisco UCS Software](#).

Other Documentation Resources

Follow [Cisco UCS Docs on Twitter](#) to receive document update notifications.

Obtaining Documentation and Submitting a Service Request

For information on obtaining documentation, submitting a service request, and gathering additional information, see the monthly [What's New in Cisco Product Documentation](#), which also lists all new and revised Cisco technical documentation.

Subscribe to the *What's New in Cisco Product Documentation* as a Really Simple Syndication (RSS) feed and set content to be delivered directly to your desktop using a reader application. The RSS feeds are a free service and Cisco currently supports RSS version 2.0.

Follow [Cisco UCS Docs on Twitter](#) to receive document update notifications.



CHAPTER 1

Overview

This chapter contains the following sections:

- [Cisco UCS B480 M5 Blade Server, on page 1](#)
- [External Features Overview, on page 3](#)
- [Drive Bays, on page 5](#)
- [Graphics Processing Unit, on page 5](#)

Cisco UCS B480 M5 Blade Server

The Cisco UCS B480 M5 is a density-optimized, full-width blade server that supports four CPU sockets for the Intel Xeon Processor Scalable Family of CPUs. The server supports the following features:

- 48 DDR4 DIMMs
- 2 front mezzanine modules (storage or graphics processing unit (GPU))
- 1 modular LAN on motherboard (mLOM) module
- 2 rear mezzanine modules (I/O or GPU)
- A mini-storage module socket with these options:
 - SD card module with two SD card slots
 - M.2 module with slots for two SATA M.2 drives
 - Cisco Boot-Optimized M.2 RAID Controller (module with two slots for SATA M.2 drives, plus an integrated SATA RAID controller that can control the two M.2 drives in a RAID 1 array)

You can install up to four UCS B480 M5 blade servers in a UCS 5108 chassis, mixing with other models of Cisco UCS blade servers in the chassis if desired.

Figure 1: Cisco UCS B480 M5 Blade Server Front Panel

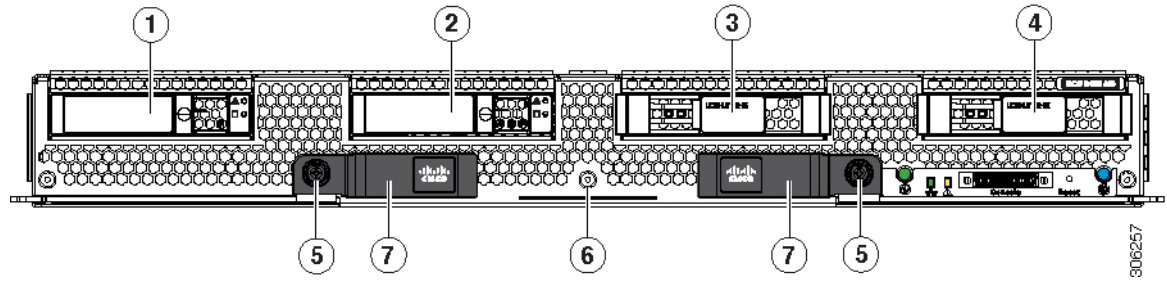
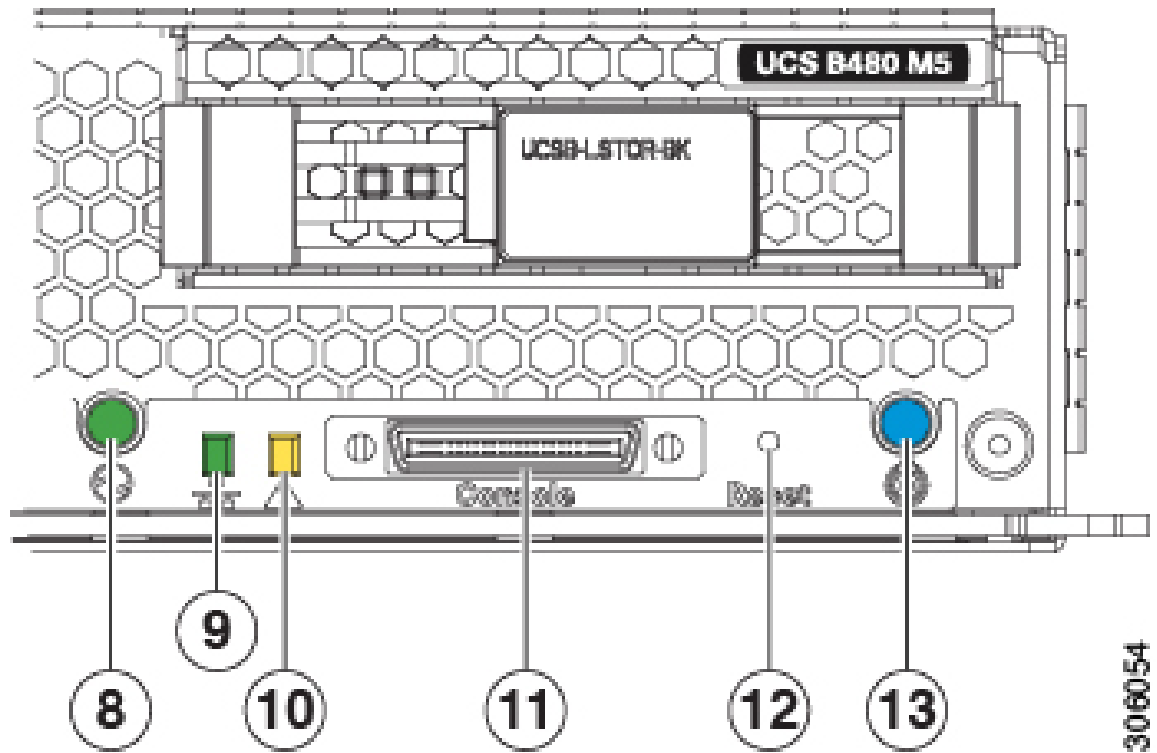


Figure 2: Cisco UCS B480 M5 Blade Server Front Panel LEDs



1	Drive bay 1	2	Drive bay 2
3	Drive bay 3	4	Drive bay 4
5	Ejector thumb screw	6	Asset pull tag
7	Blade ejector handle	8	Power button and LED
9	Network link status LED	10	Blade health LED
11	Local console connection	12	Reset button
13	Locate button and LED	–	

External Features Overview

The features of the blade server that are externally accessible are described in this section.


LEDs

Server LEDs indicate whether the blade server is in active or standby mode, the status of the network link, the overall health of the blade server, and whether the server is set to give a blinking blue locator light from the locator button.

The removable drives also have LEDs indicating hard disk access activity and disk health.

Table 1: Blade Server LEDs

LED	Color	Description
Power	Off	Power off.
	Green	Main power state. Power is supplied to all server components and the server is operating normally.
	Amber	Standby power state. Power is supplied only to the service processor of the server so that the server can still be managed. Note The front-panel power button is disabled by default. It can be re-enabled through Cisco UCS Manager. After it's enabled, if you press and release the front-panel power button, the server performs an orderly shutdown of the 12 V main power and goes to standby power state. You cannot shut down standby power from the front-panel power button. See the Cisco UCS Manager Configuration Guides for information about completely powering off the server from the software interface.

LED	Color	Description
Link	Off	None of the network links are up.
	Green	At least one network link is up.
Health	Off	Power off.
	Green	Normal operation.
	Amber	Minor error.
	Blinking Amber	Critical error.
Blue locator button and LED	Off	Blinking is not enabled.
	Blinking blue 1 Hz	Blinking to locate a selected blade—If the LED is not blinking, the blade is not selected. You can control the blinking in UCS Manager or by using the blue locator button/LED.
Activity (Disk Drive) 	Off	Inactive.
	Green	Outstanding I/O to disk drive.
Health (Disk Drive)	Off	Can mean either no fault detected or the drive is not installed.
	Flashing Amber 4 hz	Rebuild drive active. If the Activity LED is also flashing amber, a drive rebuild is in progress.
	Amber	Fault detected.

Buttons

The Reset button is recessed in the front panel of the server. You can press the button with the tip of a paper clip or a similar item. Hold the button down for five seconds, and then release it to restart the server if other methods of restarting do not work.

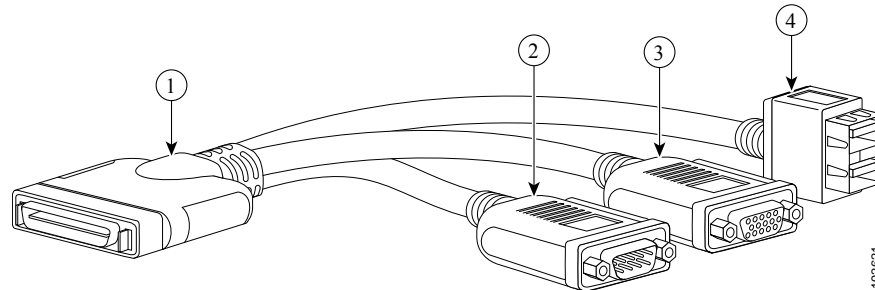
The locator function for an individual server may get turned on or off by pressing the locator button/LED.

The front-panel power button is disabled by default. It can re-enabled through Cisco UCS Manager. After it's enabled, The power button allows you to manually take a server temporarily out of service but leave it in a state where it can be restarted quickly. If the desired power state for a service profile associated with a blade server is set to "off," using the power button or Cisco UCS Manager to reset the server will cause the desired power state of the server to become out of sync with the actual power state and the server may unexpectedly shut down at a later time. To safely reboot a server from a power-down state, use the Boot Server action in Cisco UCS Manager.

Local Console Connection

The local console connector allows a direct connection to a blade server to allow operating system installation and other management tasks to be done directly rather than remotely. The port uses the KVM dongle cable that provides a connection into a Cisco UCS blade server; it has a DB9 serial connector, a VGA connector for a monitor, and dual USB ports for a keyboard and mouse. With this cable, you can create a direct connection to the operating system and the BIOS running on a blade server. A KVM cable ships standard with each blade chassis accessory kit.

Figure 3: KVM Cable for Blade Servers



1	Connector to blade server local console connection	2	DB9 serial connector
3	VGA connector for a monitor	4	2-port USB connector for a mouse and keyboard

Drive Bays

The Cisco UCS B480 M5 blade server has two front mezzanine slots that can support either storage modules or graphics processing units (GPUs). The storage module has two drive bays that can be configured with any combination of 2.5-inch SAS, SATA, or PCIe drives. A blanking panel (UCSB-LSTOR-BK) must cover all empty drive bays.

Graphics Processing Unit

An NVIDIA GPU can be installed in the front mezzanine slot of the server. When a GPU is installed, the drive bay is not present. Two blanking panels (UCSB-LSTOR-BK) are required when the GPU is installed in the front mezzanine slot. For additional information about the GPU, see [NVIDIA P6 Graphics Processing Unit, on page 46](#).



CHAPTER 2

Installing a Blade Server

This chapter contains the following sections:

- [Installing a Blade Server, on page 7](#)
- [Server Configuration, on page 9](#)
- [Powering Off a Blade Server Using the Power Button, on page 9](#)
- [Removing a Blade Server, on page 10](#)
- [Server Troubleshooting, on page 10](#)

Installing a Blade Server

For installations of UCS blades with differing widths and heights in a chassis, the guideline is to load the heaviest and largest blades at the bottom of the chassis. Therefore, if a UCS B460 blade server is present, it should be installed at the bottom, followed by full-width blades such as the UCS B420 or UCS B480 above the UCS B460, and then half-width blades such as the UCS B200 at the top of the chassis.

Before you begin

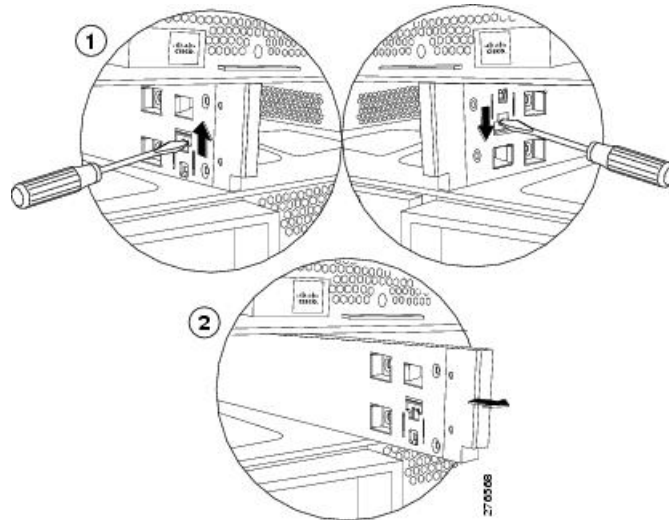
The blade server must have its cover installed before installing it into the chassis to ensure adequate airflow.

Procedure

Step 1

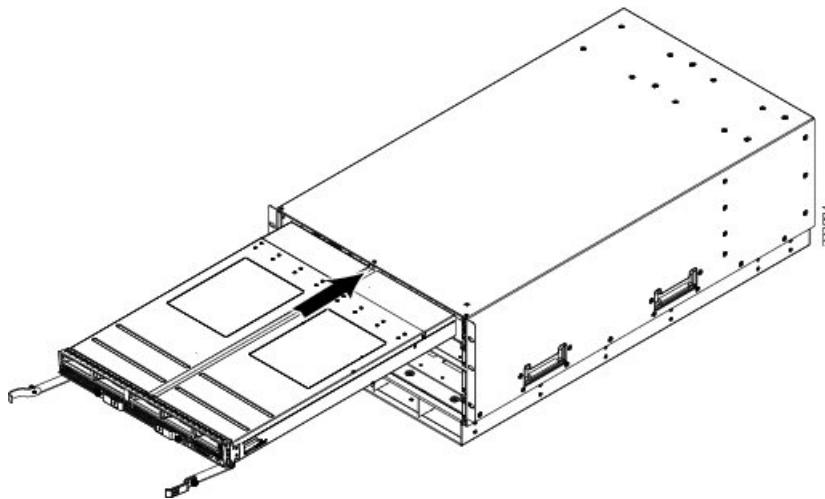
If necessary, remove the slot divider from the chassis.

- Simultaneously pull up on the left side catch and push down on the right side catch as shown in callout 1 of the following figure.
- Pull the slot divider out of the chassis as shown in callout 2 of the following figure. Keep the slot divider in case it is needed at another time.

Figure 4: Removing a Slot Divider

Tip To reinstall the slot divider, align it with the dimples in the slot top and bottom and slide it back in until it clicks into place.

Step 2 Grasp the front of the blade server and place your other hand under the blade to support it.

Figure 5: Positioning a Blade Server in the Chassis

Step 3 Open the ejector levers in the front of the blade server.

Step 4 Gently slide the blade into the opening until you cannot push it any farther.

Step 5 Press the ejectors so that they catch the edge of the chassis and press the blade server all the way in.

Step 6 Tighten the captive screw on the front of the blade to no more than 3 in-lbs. Tightening only with bare fingers is unlikely to lead to stripped or damaged captive screws.

Cisco UCS Manager automatically reacknowledges, reassociates, and recommissions the server, provided any hardware changes are allowed by the service profile.

Server Configuration

Cisco UCS blade servers should be configured and managed using Cisco UCS Manager. For details, see the *Configuration Guide* for the version of Cisco UCS Manager that you are using. The configuration guides are available at the following URL:

http://www.cisco.com/en/US/products/ps10281/products_installation_and_configuration_guides_list.html

Powering Off a Blade Server Using the Power Button



Note The front panel power button is disabled by default to ensure that servers are decommissioned through the UCS Manager software before shutdown. If you prefer to shut down the server locally with the button, you can enable front power-button control in UCS Manager.



Tip You can also shut down servers remotely using Cisco UCS Manager. For details, see the *Configuration Guide* for the version of Cisco UCS Manager that you are using. The configuration guides are available at the following URL:

http://www.cisco.com/en/US/products/ps10281/products_installation_and_configuration_guides_list.html

Procedure

- Step 1** If you are local to the server, check the color of the **Power Status** LED for each server in the chassis that you want to power off.
- Green indicates that the server is running and must be shut down before it can be safely powered off. Go to Step 2.
 - Amber indicates that the server is already in standby mode and can be safely powered off. Go to Step 3.
- Step 2** If you previously enabled front power-button control through Cisco UCS Manager, press and release the **Power** button, then wait until the **Power Status** LED changes to amber.
- The operating system performs a graceful shutdown and the server goes to standby mode.
- Caution** To avoid data loss or damage to your operating system, you should always invoke a graceful shutdown of the operating system.
- Step 3** (Optional) Although not recommended, if you are shutting down all blade servers in a chassis, you can disconnect the power cords from the chassis to completely power off the servers.

Caution To avoid data loss or damage to your operating system, you should always invoke a graceful shutdown of the operating system.

The blade servers will power down. You can now perform additional tasks with the blades as needed, for example, replacing a blade.

Removing a Blade Server

Using UCS Manager, decommission the server before physically removing the server. To remove a blade server from the chassis, follow these steps:

Procedure

- Step 1** Loosen the captive screws on the front of the blade.
 - Step 2** Remove the blade from the chassis by pulling the ejector lever on the blade until it unseats the blade server.
 - Step 3** Slide the blade partially out of the chassis, and place your other hand under the blade to support its weight.
 - Step 4** Once completely removed, place the blade on an antistatic mat or antistatic foam if you are not immediately reinstalling it into another slot.
 - Step 5** If the slot is to remain empty, install a blank faceplate (N20-CBLKB1) to maintain proper thermal temperature and keep dust out of the chassis.
-

Server Troubleshooting

For general troubleshooting information, see the [Cisco UCS Manager Troubleshooting Reference Guide](#).



CHAPTER 3

Servicing a Blade Server

This chapter contains the following sections:

- [Replacing a Drive, on page 11](#)
- [Removing a Blade Server Cover, on page 14](#)
- [Internal Components, on page 14](#)
- [Diagnostics Button and LEDs, on page 16](#)
- [Installing the Front Mezzanine Storage Module, on page 16](#)
- [Replacing the SuperCap Module, on page 17](#)
- [Replacing CPUs and Heatsinks, on page 20](#)
- [Replacing Memory DIMMs, on page 35](#)
- [Replacing Intel Optane DC Persistent Memory Modules, on page 39](#)
- [Installing a Virtual Interface Card in the mLOM Slot, on page 44](#)
- [Installing a Rear Mezzanine Module in Addition to the mLOM VIC, on page 45](#)
- [NVIDIA P6 Graphics Processing Unit, on page 46](#)
- [Enabling the Trusted Platform Module, on page 52](#)
- [Removing the Trusted Platform Module \(TPM\), on page 54](#)
- [Mini Storage Module, on page 55](#)
- [Replacing a Boot-Optimized M.2 RAID Controller Module, on page 67](#)
- [Recycling the PCB Assembly \(PCBA\), on page 71](#)

Replacing a Drive

The Cisco UCS B480 M5 blade server can use two optional front storage mezzanine modules that each have two drive bays for hard disks or SSD drives, either 2.5-inch SAS, SATA, or NVMe. The storage mezzanine module also supports a RAID controller. If you purchased the server without a front storage mezzanine module configured as a part of the system, a pair of blanking panels may be in place. These panels should be removed before installing disk drives, but should remain in place to ensure proper cooling and ventilation if the drive bays are unused.

You can remove and install disk drives without removing the blade server from the chassis.



Caution

To avoid data loss or damage to your operating system, always perform drive service during a scheduled maintenance window.

The drives supported in this blade server come with the hot-plug drive sled attached. Empty hot-plug drive sled carriers (containing no drives) are not sold separately from the drives. A list of currently supported drives is in the *Cisco UCS B480 M5 Blade Server Data Sheet* on the [Cisco UCS B-Series Blade Servers Data Sheets](#) page.

Before upgrading or adding a drive to a running blade server, check the service profile in Cisco UCS Manager and make sure the new hardware configuration will be within the parameters allowed by the service profile.



Note See also [4K Sector Format SAS/SATA Drives Considerations](#), on page 13.

Removing a Blade Server Drive

To remove a drive from a blade server, follow these steps:

Procedure

- Step 1** Push the release button to open the drive ejector, and then pull the drive from its slot.
- Caution** To prevent data loss, make sure that you know the state of the system before removing a drive.
- Step 2** Place the drive on an antistatic mat or antistatic foam if you are not immediately reinstalling it in another server.
- Step 3** Install a drive blanking panel to maintain proper airflow and keep dust out of the drive bay if it will remain empty.
-

Installing a Blade Server Drive

To install a drive in a blade server, follow these steps:

Procedure

- Step 1** Place the drive ejector into the open position by pushing the release button.
- Step 2** Gently slide the drive into the opening in the blade server until it seats into place.
- Step 3** Push the drive ejector into the closed position.

You can use Cisco UCS Manager to format and configure RAID services. For details, see the *Configuration Guide* for the version of Cisco UCS Manager that you are using. The configuration guides are available at the following URL:

http://www.cisco.com/en/US/products/ps10281/products_installation_and_configuration_guides_list.html

If you need to move a RAID cluster, see the [Cisco UCS Manager Troubleshooting Reference Guide](#).

4K Sector Format SAS/SATA Drives Considerations

- You must boot 4K sector format drives in UEFI mode, not legacy mode. See the procedure in this section for setting UEFI boot mode in the boot policy.
- Do not configure 4K sector format and 512-byte sector format drives as part of the same RAID volume.
- For operating system support on 4K sector drives, see the interoperability matrix tool for your server: [Hardware and Software Interoperability Matrix Tools](#)

Setting Up UEFI Mode Booting in the UCS Manager Boot Policy

Procedure

- Step 1** In the Navigation pane, click **Servers**.
- Step 2** Expand **Servers > Policies**.
- Step 3** Expand the node for the organization where you want to create the policy.
If the system does not include multitenancy, expand the root node.
- Step 4** Right-click **Boot Policies** and select **Create Boot Policy**.
The Create Boot Policy wizard displays.
- Step 5** Enter a unique name and description for the policy.
This name can be between 1 and 16 alphanumeric characters. You cannot use spaces or any special characters other than - (hyphen), _ (underscore), : (colon), and . (period). You cannot change this name after the object is saved.
- Step 6** (Optional) After you make changes to the boot order, check the **Reboot on Boot Order Change** check box to reboot all servers that use this boot policy.
For boot policies applied to a server with a non-Cisco VIC adapter, even if the Reboot on Boot Order Change check box is not checked, when SAN devices are added, deleted, or their order is changed, the server always reboots when boot policy changes are saved.
- Step 7** (Optional) If desired, check the **Enforce vNIC/vHBA/iSCSI Name** check box.
- If checked, Cisco UCS Manager displays a configuration error and reports whether one or more of the vNICs, vHBAs, or iSCSI vNICs listed in the Boot Order table match the server configuration in the service profile.
 - If not checked, Cisco UCS Manager uses the vNICs or vHBAs (as appropriate for the boot option) from the service profile.
- Step 8** In the **Boot Mode** field, choose the **UEFI** radio button.
- Step 9** Check the Boot Security check box if you want to enable UEFI boot security.
- Step 10** Configure one or more of the following boot options for the boot policy and set their boot order:
- Local Devices boot—To boot from local devices, such as local disks on the server, virtual media, or remote virtual disks, continue with *Configuring a Local Disk Boot for a Boot Policy* in the [Cisco UCS Manager Server Management Guide](#) for your release.

- SAN boot—To boot from an operating system image on the SAN, continue with *Configuring a SAN Boot for a Boot Policy* in the [Cisco UCS Manager Server Management Guide](#) for your release.

You can specify a primary and a secondary SAN boot. If the primary boot fails, the server attempts to boot from the secondary.

- LAN boot—To boot from a centralized provisioning server, continue with *Configuring a LAN Boot For a Boot Policy* in the [Cisco UCS Manager Server Management Guide](#) for your release.
- iSCSI boot—To boot from an iSCSI LUN, continue with *Creating an iSCSI Boot Policy* in the [Cisco UCS Manager Server Management Guide](#) for your release.

Removing a Blade Server Cover

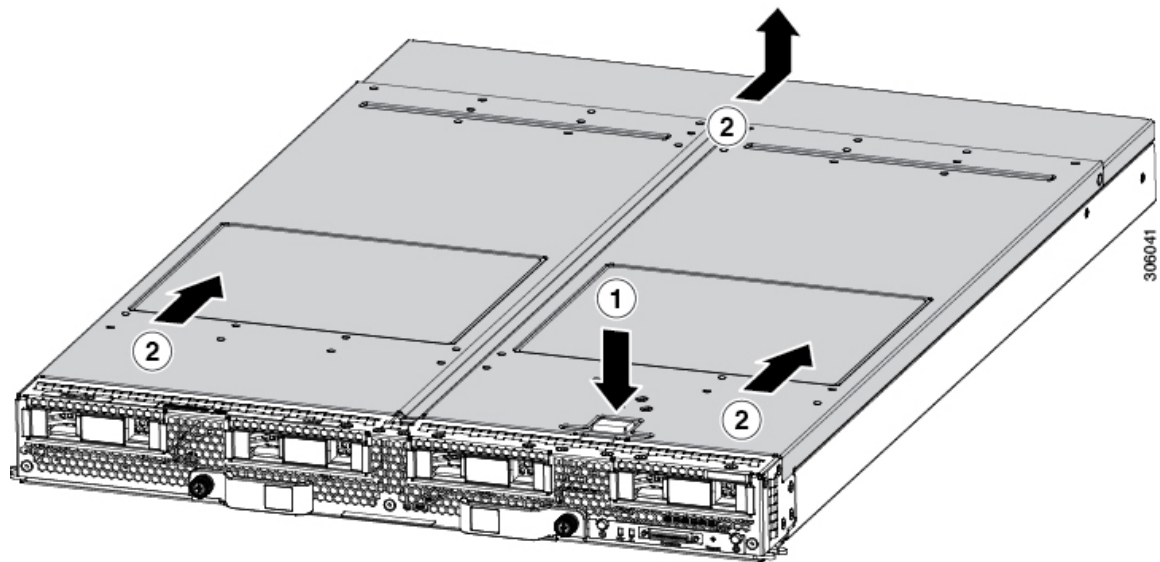
To remove the cover of the blade server, follow these steps:

Procedure

Step 1 Press and hold the button down as shown in the figure below.

Step 2 While holding the back end of the cover, pull the cover back and then up.

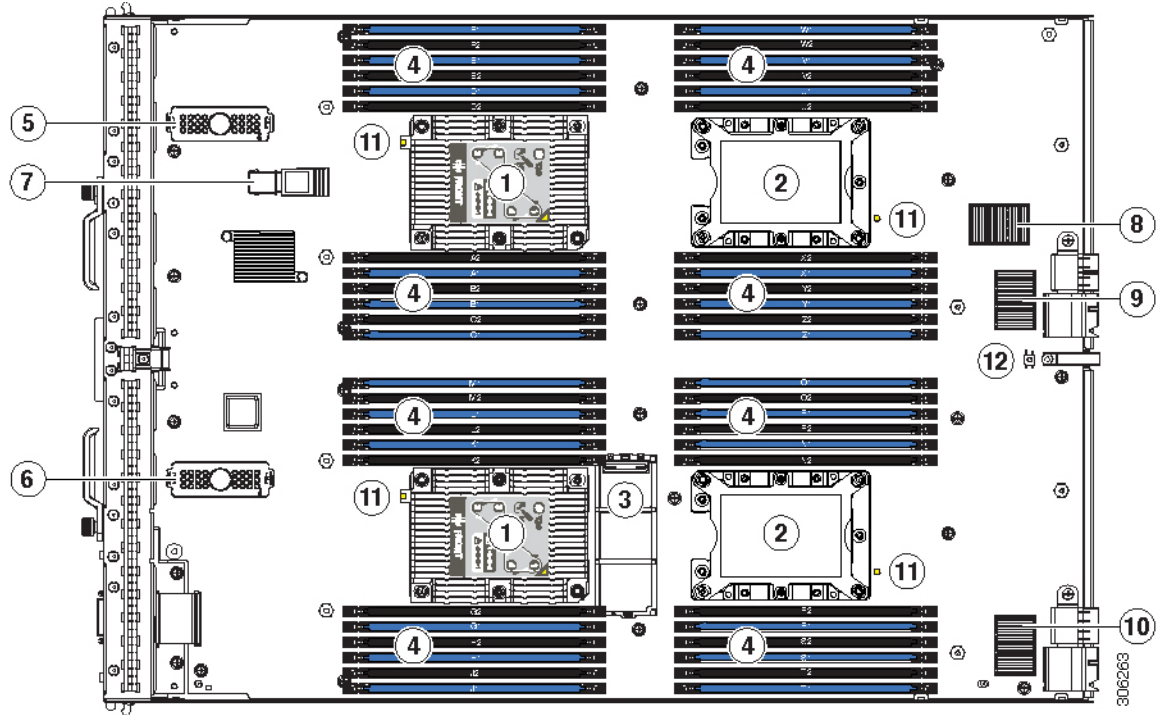
Figure 6: Removing the Cover of the Blade Server



Internal Components

The following figure shows the internal components of the Cisco UCS B480 M5 blade server.

Figure 7: Inside View of the Cisco UCS B480 M5 Blade Server



1	CPU 1 socket (populated, left side) and CPU 2 socket (populated, right side)	2	CPU 3 socket (right side) and CPU 4 socket (left side)
3	Mini storage connector	4	DIMM slots
5	Front mezzanine connector 1	6	Front mezzanine connector 2
7	USB connector (populated) An internal USB 3.0 port is supported. A 16 GB USB drive (UCS-USBFLSHB-16GB) is available from Cisco. A clearance of 0.950 inches (24.1 mm) is required for the USB device to be inserted and removed.	8	mLOM connector
9	Rear mezzanine connector 1	10	Rear mezzanine connector 2
11	CPU heat sink install guide pins	12	Diagnostic button



Note When the front mezzanine storage module is installed, the USB connector is underneath it. Use the small cutout opening in the storage module to visually determine the location of the USB connector when you need to insert a USB drive. When the NVIDIA GPU is installed in the front mezzanine slot, you cannot see the USB connector.

Diagnostics Button and LEDs

At blade start-up, POST diagnostics test the CPUs, DIMMs, HDDs, and rear mezzanine modules, and any failure notifications are sent to Cisco UCS Manager. You can view these notifications in the Cisco UCS Manager System Error Log or in the output of the **show tech-support** command. If errors are found, an amber diagnostic LED also lights up next to the failed component. During run time, the blade BIOS and component drivers monitor for hardware faults and will light up the amber diagnostic LED as needed.

LED states are saved, and if you remove the blade from the chassis the LED values will persist for up to 10 minutes. Pressing the LED diagnostics button on the motherboard causes the LEDs that currently show a component fault to light for up to 30 seconds for easier component identification. LED fault values are reset when the blade is reinserted into the chassis and booted, and the process begins from its start.

If DIMM insertion errors are detected, they may cause the blade discovery process to fail and errors will be reported in the server POST information, which is viewable using the UCS Manager GUI or CLI. DIMMs must be populated according to specific rules. The rules depend on the blade server model. Refer to the documentation for a specific blade server for those rules.

Faults on the DIMMs or rear mezzanine modules also cause the server health LED to light solid amber for minor error conditions or blinking amber for critical error conditions.

Installing the Front Mezzanine Storage Module

The Cisco UCS B480 M5 blade server can use two optional front storage mezzanine modules that each have two drive bays and RAID controller or NVMe-based PCIe SSD support functionality.



Note There are restrictions against using the front mezzanine storage module with certain CPUs because of heat concerns. See [CPU Configuration Rules, on page 21](#).

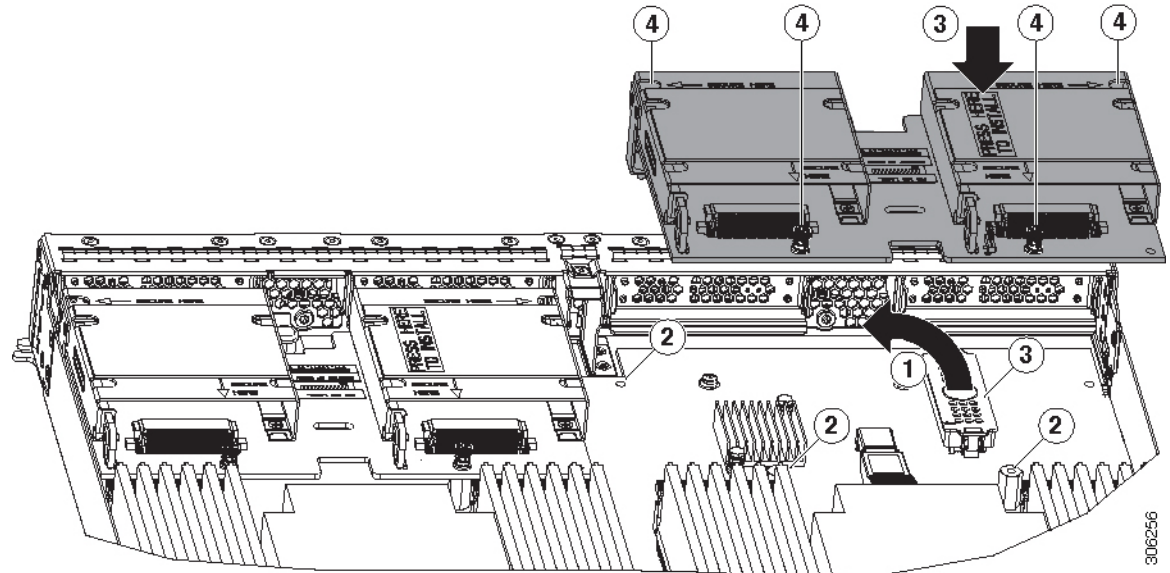
To install the front mezzanine storage module, follow these steps:

Procedure

- Step 1** Remove the connectors' protective covers (callout 1) from both the front mezzanine storage module and the motherboard.
- Step 2** Place the storage module over the front mezzanine connector and the two standoff posts (callout 2) on the motherboard at the front of the server.

- Step 3** Press down on the drive bay cage where it is labeled "Press Here to Install" (callout 3) until the storage module clicks into place.

Figure 8: Front Mezzanine Storage Module



- Step 4** Using a Phillips-head screwdriver, tighten the four screws (callout 4) to secure the storage module. The locations of the screws are labeled "Secure Here."

Replacing the SuperCap Module

The SuperCap module is a battery bank which connects to the front mezzanine storage module board and provides power to the RAID controller if facility power is interrupted.

To replace the SuperCap module, use the following topics:

- [Removing the SuperCap Module, on page 17](#)
- [Installing the SuperCap Module, on page 19](#)

Removing the SuperCap Module

The SuperCap module sits in a plastic tray. The module connects to the board through a ribbon cable with one connector to the module and one connector to the board. The SuperCap replacement PID (UCSB-MRAID-SC=) contains the module only, so you must leave the ribbon cable in place on the board.



- Caution** When disconnecting the SuperCap module, disconnect the ribbon cable from the module only. Do not disconnect the cable from the board. The board connection and the tape that secures the cable must remain connected and undamaged.

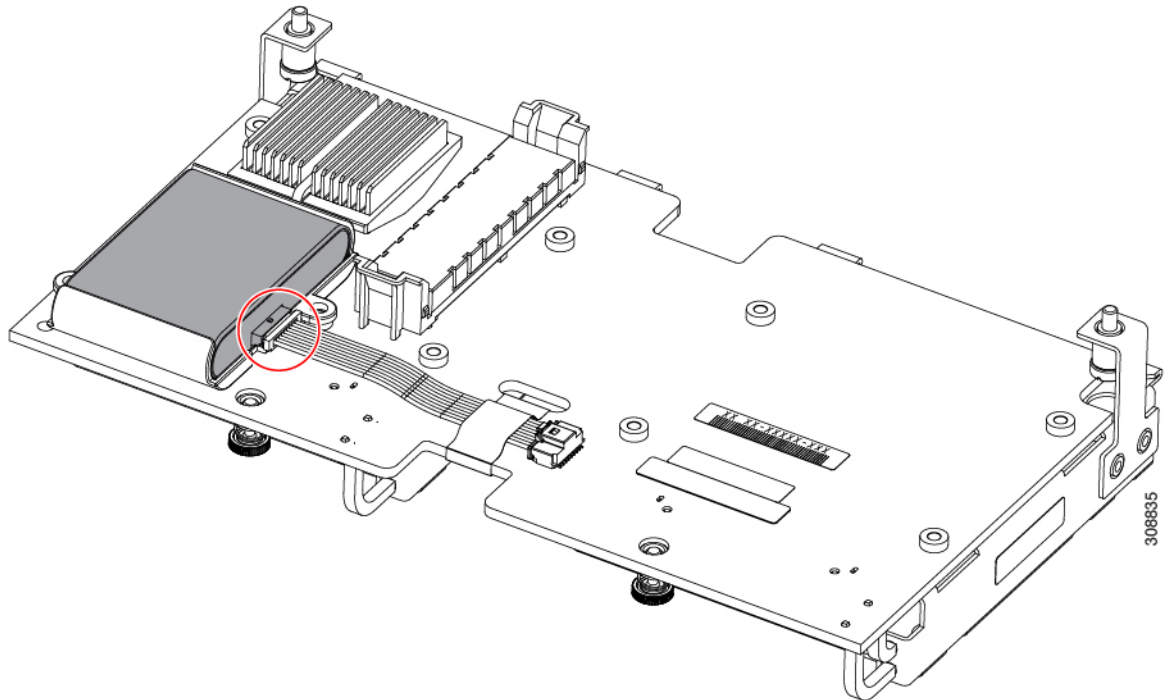
To replace the SuperCap module, follow these steps:

Procedure

Step 1 Grasp the cable connector at the SuperCap module and gently pull to disconnect the cable from the SuperCap module.

Do not grasp the cable itself, the tape, or the board connector.

Figure 9: Disconnecting the SuperCap Cable from the Module, Not the Board

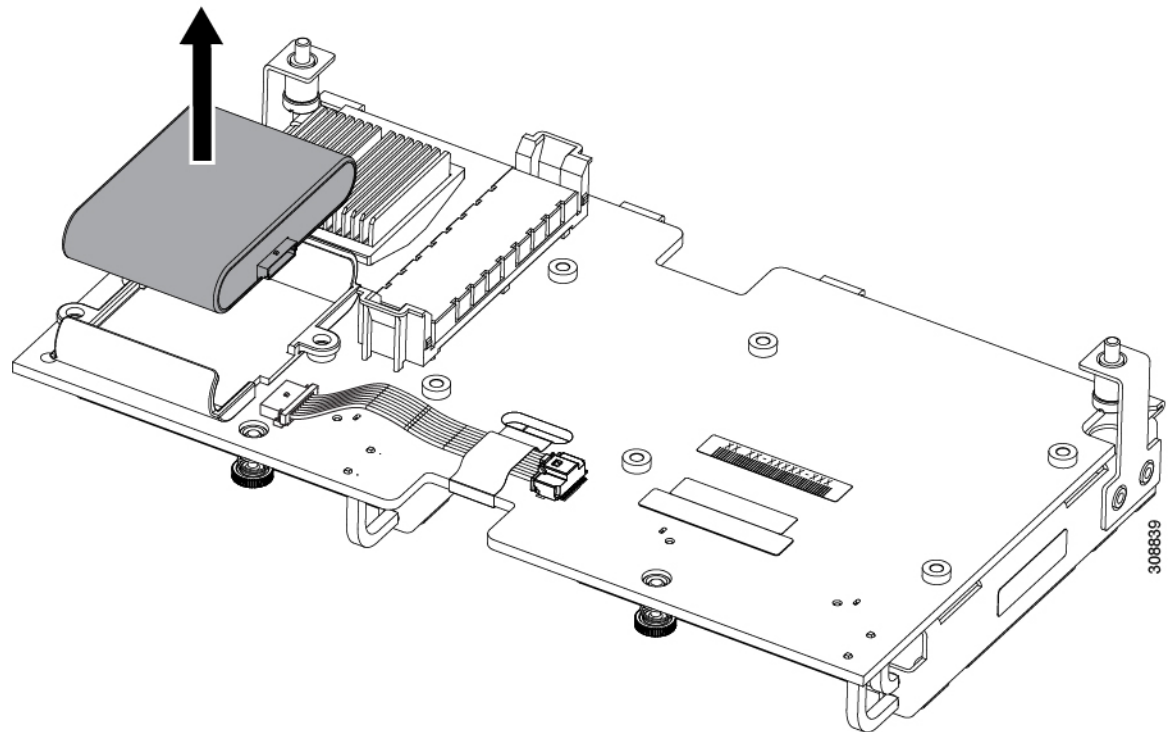


Step 2 Before removing the SuperCap module, note its orientation in the tray.

When correctly oriented, the connector is on the bottom half of the module and faces the cable. You will need to install the new SuperCap module with the same orientation.

Step 3 Grasp the sides of the SuperCap module, but not the connector, and lift the SuperCap module out of the tray.

Figure 10: Removing the SuperCap Module



You might feel some resistance because the tray is curved to secure the module.

Installing the SuperCap Module

To install a SuperCap module (UCSB-MRAID-SC=), use the following steps:

Procedure

Step 1 Orient the SuperCap module correctly, as shown (1).

When correctly oriented:

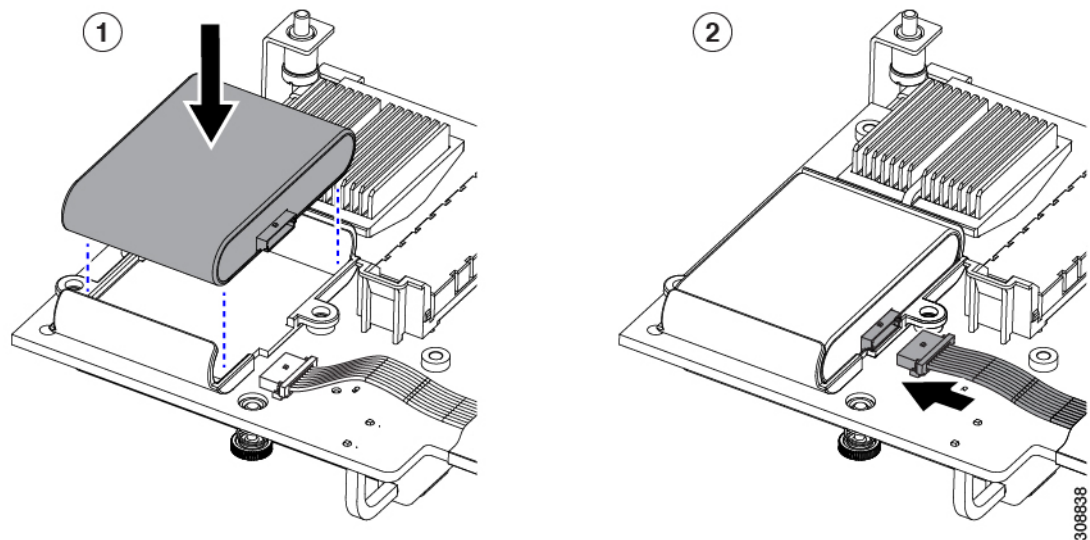
- The connector is on the bottom half of the module facing the cable.
- The connector will fit into the rectangular notch in the tray. This notch is specifically designed to accept the SuperCap module connector.

Caution Make sure the SuperCap module is properly oriented before proceeding. If the module is installed incorrectly, the ribbon cable can get snagged or damaged.

Step 2 When the module is correctly oriented, lower the module and press down until it clips into the tray.

You might feel some resistance while the module passes the curved clips at the top of the tray.

Figure 11: Orienting and Installing the SuperCap Module

**Step 3**

When the module is seated in the tray, reconnect the cable (2):

- a) Grasp the cable connector and verify that the pins and sockets on the cable connector and module connector are correctly aligned.
- b) When the cable connector and module connector are properly aligned, plug the cable into the SuperCap module.

What to do next

Reinstall the blade server. Go to [Installing a Blade Server](#), on page 7.

Replacing CPUs and Heatsinks

This topic describes the configuration rules and procedure for replacing CPUs and heatsinks.

Special Information For *Upgrades* to Second Generation Intel Xeon Scalable Processors

**Caution**

You must upgrade your server firmware to the required minimum level before you upgrade to the Second Generation Intel Xeon Scalable processors that are supported in this server. Older firmware versions cannot recognize the new CPUs and this would result in a non-bootable server.

The minimum software and firmware versions required for this server to support Second Generation Intel Xeon Scalable processors are as follows:

Table 2: Minimum Requirements For Second Generation Intel Xeon Scalable Processors

Software or Firmware	Minimum Version
Cisco UCS Manager	4.0(4)
Server Cisco IMC	4.0(4)
Server BIOS	4.0(4)

Do one of the following actions:

- If your server's firmware and Cisco UCS Manager software are already at the required minimums shown above (or later), you can replace the CPU hardware by using the procedure in this section.
- If your server's firmware and Cisco UCS Manager software are earlier than the required levels, use the instructions in the [Cisco UCS B-Series M5 Servers Upgrade Guide For Next Gen Intel Xeon Processors](#) to upgrade your software. After you upgrade the software, return to this section as directed to replace the CPU hardware.

CPU Configuration Rules

This server has four CPU sockets on the motherboard. Each CPU supports six DIMM channels (12 DIMM slots). See [Memory Population Guidelines, on page 35](#).

- The server can operate with two or four identical CPUs installed.
- The minimum configuration is at least CPU 1 and CPU 2 installed. Install CPUs 1 and 2 first and then CPUs 3 and 4.

The following restrictions apply when using a dual-CPU configuration:

- Any unused CPU socket must have the protective dust cover from the factory installed.
- The maximum number of DIMMs is 24 (installed in slots A through J).
- Mezzanine slots 1 and 2 are unavailable.
- **For Intel Xeon Scalable processors (first generation):** The maximum combined memory allowed in the 12 DIMM slots controlled by any one CPU is 768 GB. To populate the 12 DIMM slots with more than 768 GB of combined memory, you must use a high-memory CPU that has a PID that ends with an "M", for example, UCS-CPU-6134M.
- **For Second Generation Intel Xeon Scalable processors:** These Second Generation CPUs have three memory tiers. These rules apply on a *per-socket* basis:
 - If the CPU socket has up to 1 TB of memory installed, a CPU with no suffix can be used (for example, Gold 6240).
 - If the CPU socket has 1 TB or more (up to 2 TB) of memory installed, you must use a CPU with an M suffix (for example, Platinum 8276M).
 - If the CPU socket has 2 TB or more (up to 4.5 TB) of memory installed, you must use a CPU with an L suffix (for example, Platinum 8270L).

**Caution**

In the following table, systems configured with the processors shown must adhere to the ambient inlet temperature thresholds specified. If not, a fan fault or executing workloads with extensive use of heavy instructions sets such as Intel® Advanced Vector Extensions 512 (Intel® AVX-512) may assert thermal and/or performance faults with an associated event recorded in the System Event Log (SEL). The table lists ambient temperature limitations below 35° C (95° F) and configuration restrictions to ensure proper cooling and avoid excessive processor throttling, which may impact system performance.

Table 3: Ambient Temperature and Configuration Restrictions

Processor Thermal Design Power (TDP)	CPU PID	Blade Slot	Ambient Temperature Limitation	Configuration Restriction
Any Y or N SKUs	UCS-CPU-18260Y UCS-CPU-16240Y UCS-CPU-16230N	Any	32° C [90° F]	Front Mezzanine GPU
200W or 205W	UCS-CPU-18280M UCS-CPU-18280L UCS-CPU-18280 UCS-CPU-18270 UCS-CPU-18268 UCS-CPU-8180M UCS-CPU-8180 UCS-CPU-8168 UCS-CPU-16254 UCS-CPU-6154	Any		
Frequency Optimized 150/165/125W	UCS-CPU-16246 UCS-CPU-16244 UCS-CPU-15222	Any		

Tools Required for CPU Replacement

You need the following tools and equipment for this procedure:

- T-30 Torx driver—Supplied with replacement CPU.
- #1 flat-head screwdriver—Supplied with replacement CPU.
- CPU assembly tool—Supplied with replacement CPU. Can be ordered separately as Cisco PID UCS-CPUAT=.
- Heatsink cleaning kit—Supplied with replacement CPU. Can be ordered separately as Cisco PID UCSX-HSCK=.

One cleaning kit can clean up to four CPUs.

- Thermal interface material (TIM)—Syringe supplied with replacement CPU. Use only if you are reusing your existing heatsink (new heatsinks have pre-applied TIM). Can be ordered separately as Cisco PID UCS-CPU-TIM=.

One TIM kit covers one CPU.

Replacing a CPU and Heatsink

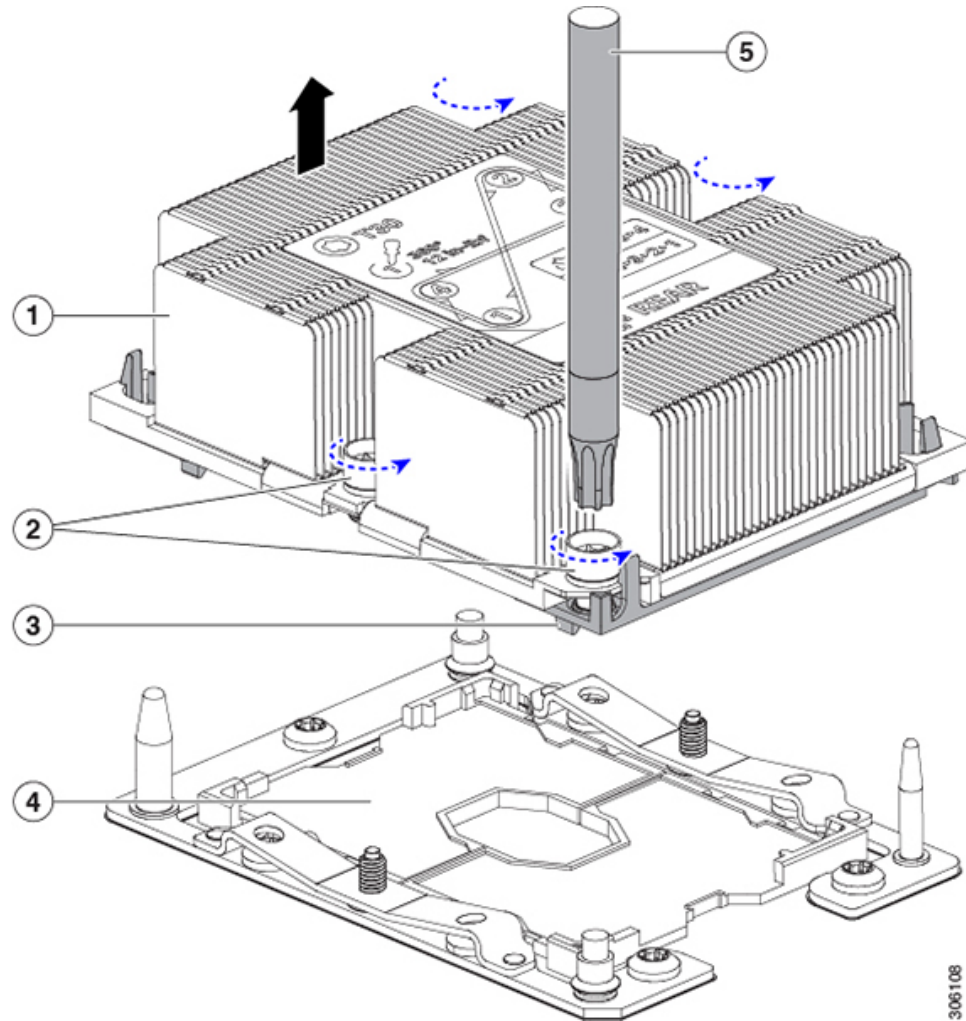


Caution CPUs and their sockets are fragile and must be handled with extreme care to avoid damaging pins. The CPUs must be installed with heatsinks and thermal interface material to ensure cooling. Failure to install a CPU correctly might result in damage to the server.

Procedure

- Step 1** Remove the existing CPU/heatsink assembly from the server:
- Decommission and power off the server.
 - Slide the server out the front of the chassis.
 - Remove the top cover from the server as described in [Removing a Blade Server Cover, on page 14](#).
 - Use the T-30 Torx driver that is supplied with the replacement CPU to loosen the four captive nuts that secure the heatsink with the attached CPU assembly to the motherboard standoff.
 - Note** Alternate loosening the heatsink nuts evenly so that the heatsink remains level as it is raised. Loosen the heatsink nuts in the order shown on the heatsink label: 4, 3, 2, 1.
 - Lift straight up on the CPU/heatsink assembly and set it heatsink-down on an antistatic surface.
 - Note** Make sure to hold the heatsink along the fin edges and not the fin walls to prevent damaging the heatsink.

Figure 12: Removing the CPU/Heatsink Assembly



1	Heatsink	2	Heatsink captive nuts (two on each side)
3	CPU carrier (below heatsink in this view)	4	CPU socket on motherboard
5	T-30 Torx driver	-	

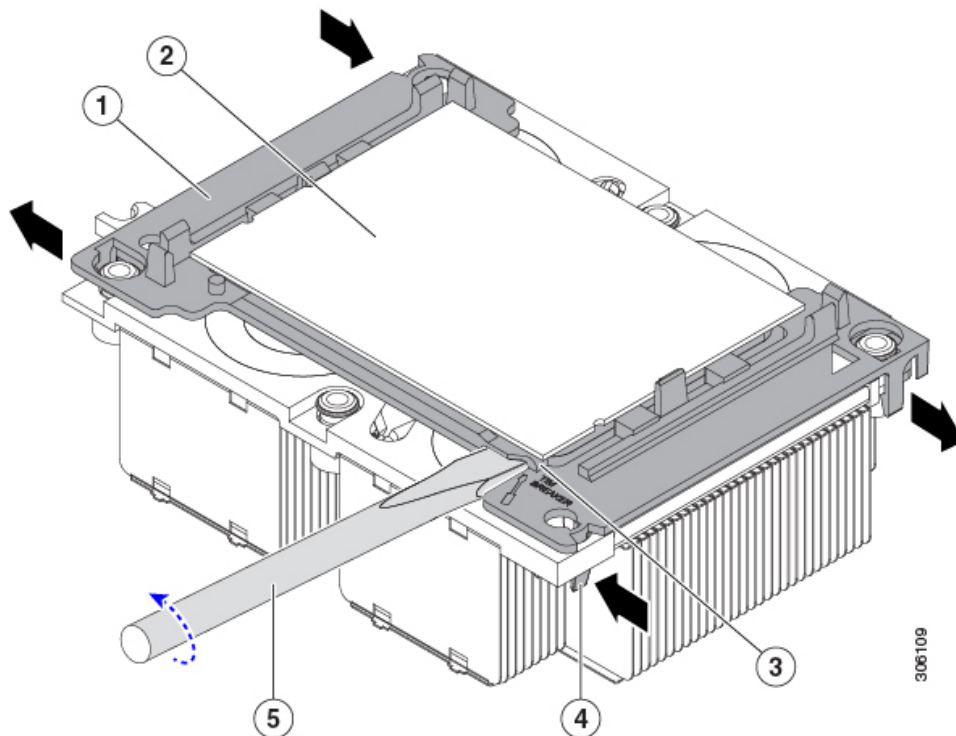
Step 2

Separate the heatsink from the CPU assembly (the CPU assembly includes the CPU and the CPU carrier):

- a) Place the heatsink with CPU assembly so that it is oriented upside-down as shown in the following figure.

Note the thermal-interface material (TIM) breaker location. TIM BREAKER is stamped on the CPU carrier next to a small slot.

Figure 13: Separating the CPU Assembly From the Heatsink



1	CPU carrier	2	CPU
3	TIM BREAKER slot in CPU carrier	4	CPU-carrier inner-latch nearest to the TIM breaker slot
5	#1 flat-head screwdriver inserted into TIM breaker slot	6	CPU carrier inner-latch at corner opposite of TIM breaker slot
7	CPU carrier outer latches	-	

- b) Pinch inward on the CPU-carrier inner-latch that is nearest the TIM breaker slot and then push up to disengage the clip from its slot in the heatsink corner.
- c) Insert the blade of a #1 flat-head screwdriver into the slot marked TIM BREAKER.
- d) Gently rotate the screwdriver to lift the CPU heat spreader until the TIM on the heatsink separates from the CPU.

Note Use caution to avoid damaging the heatsink surface. Do not allow the screwdriver tip to touch or damage the green CPU substrate.

- e) Pinch the CPU-carrier inner-latch at the corner opposite the TIM breaker and push up to disengage the clip from its slot in the heatsink corner.
- f) On the remaining two corners of the CPU carrier, gently pry outward on the outer-latches and then lift the CPU-assembly from the heatsink.

Note Handle the CPU-assembly by the plastic carrier only. Do not touch the CPU surface. Do not separate the CPU from the carrier.

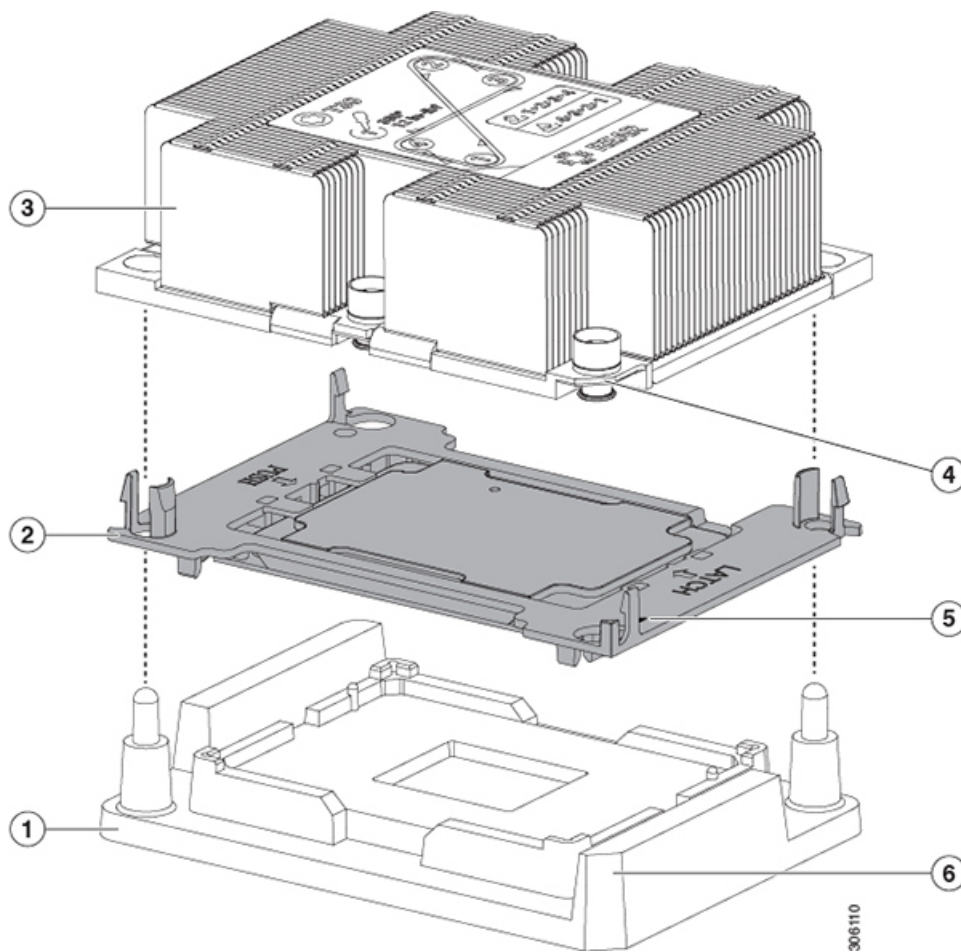
Step 3 The new CPU assembly is shipped on a CPU assembly tool. Take the new CPU assembly and CPU assembly tool out of the carton.

Caution When CPUs greater than 165 W are installed in the server, you cannot install the front mezzanine storage module because of heat concerns. There are other CPUs that require a reduction in server operating temperature (air inlet temperature). See [CPU Configuration Rules, on page 21](#).

If the CPU assembly and CPU assembly tool become separated, note the alignment features shown in the following figure for the correct orientation. The pin 1 triangle on the CPU carrier must be aligned with the angled corner on the CPU assembly tool.

Caution CPUs and their sockets are fragile and must be handled with extreme care to avoid damaging pins.

Figure 14: CPU Assembly Tool, CPU Assembly, and Heatsink Alignment Features



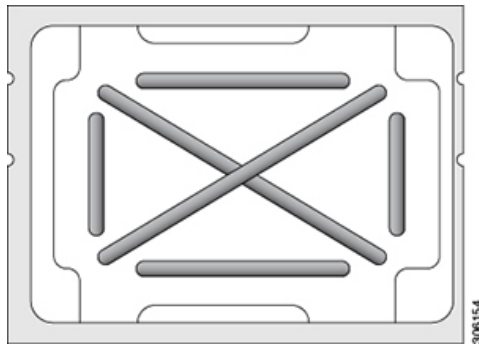
1	CPU assembly tool	2	CPU assembly (CPU in plastic carrier)
3	Heatsink	4	Angled corner on heatsink (pin 1 alignment feature)
5	Triangle cut into carrier (pin 1 alignment feature)	6	Angled corner on CPU assembly tool (pin 1 alignment feature)

Step 4 Apply new TIM.

Note The heatsink must have new TIM on the heatsink-to-CPU surface to ensure proper cooling and performance.

- If you are installing a new heatsink, it is shipped with a pre-applied pad of TIM. Go to Step 5.
 - If you are reusing a heatsink, you must remove the old TIM from the heatsink and then apply new TIM to the CPU surface from the supplied syringe. Continue with Step **a** below.
- a) Apply the Bottle #1 cleaning solution that is included with the heatsink cleaning kit (UCSX-HSCK=), as well as the spare CPU package, to the old TIM on the heatsink and let it soak for a least 15 seconds.
 - b) Wipe all of the TIM off the heatsink using the soft cloth that is included with the heatsink cleaning kit. Be careful to avoid scratching the heatsink surface.
 - c) Using the syringe of TIM provided with the new CPU (UCS-CPU-TIM=), apply 1.5 cubic centimeters (1.5 ml) of thermal interface material to the top of the CPU. Use the pattern shown in the following figure to ensure even coverage.

Figure 15: Thermal Interface Material Application Pattern



Caution Use only the correct heatsink for your CPU. CPU 1 uses heatsink UCSB-HS-M5-F and CPU 2 uses heatsink UCSB-HS-M5-R.

Step 5 With the CPU assembly on the CPU assembly tool, set the heatsink onto the CPU assembly.

- a) Place the heatsink onto the CPU by aligning the Pin 1 corner of the heatsink with the Pin 1 tab of the CPU carrier for the correct orientation.
- b) Push down gently until you hear the corner latches of the CPU carrier click onto the heatsink corners.
- c) Inspect all four latches to verify they are fully engaged.

Caution In the following step, use extreme care to avoid touching or damaging the CPU contacts or the CPU socket pins.

Step 6 Install the CPU/heatsink assembly to the server.

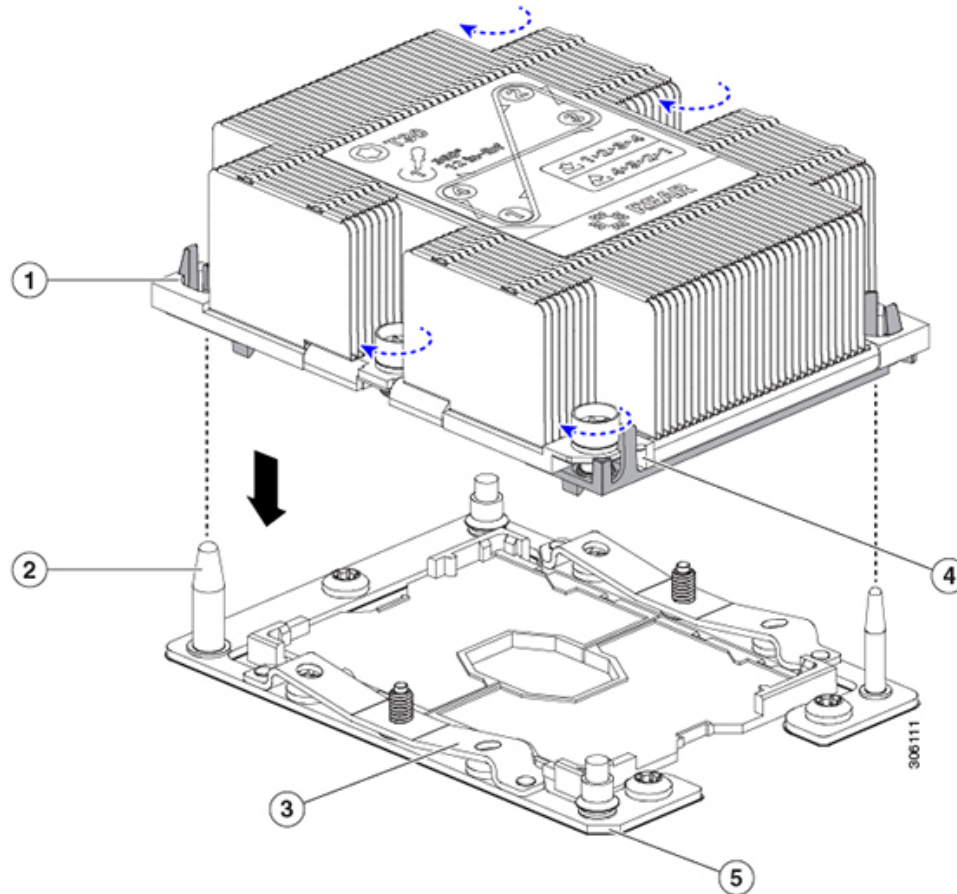
- a) Lift the heatsink with attached CPU assembly from the CPU assembly tool.

Note Make sure to hold the heatsink along the fin edges and not the fin walls to prevent damaging the heatsink.

- b) Align the CPU with heatsink over the CPU socket on the motherboard, as shown in the following figure.

Note the alignment features. The pin 1 angled corner on the heatsink must align with the pin 1 angled corner on the CPU socket. The CPU socket alignment pins must properly align with the slots on the CPU carrier and heatsink. Take note of the two different sizes of the alignment pins.

Figure 16: Installing the Heatsink/CPU Assembly to the CPU Socket



1	Guide hole in assembly (two)	2	CPU socket alignment post (two)
3	CPU socket leaf spring	4	Angled corner on heatsink (pin 1 alignment feature)
5	Angled corner on socket (pin 1 alignment feature)	6	CPU socket leaf spring threaded standoffs
7	CPU socket alignment threaded standoffs	-	

- c) Set the heatsink with CPU assembly down onto the CPU socket.
- d) Use the T-30 Torx driver that is supplied with the replacement CPU to tighten the four captive nuts that secure the heatsink to the motherboard standoffs.

Caution Alternate tightening the heatsink nuts evenly so that the heatsink remains level while it is installed. Tighten the heatsink nuts in the order shown on the heatsink label: 1, 2, 3, 4. The captive nuts must be fully tightened so that the leaf springs on the CPU socket lie flat.

- e) Replace the top cover to the server.
- f) Replace the server in the chassis.
- g) Wait for UCS Manager to complete its automatic discovery of the server.

Additional CPU-Related Parts to Order with CPU RMA

When a return material authorization (RMA) of the CPU occurs for a Cisco UCS B-Series server, additional parts might not be included with the CPU spare bill of materials (BOM). The TAC engineer might need to add the additional parts to the RMA to help ensure a successful replacement.



Note The following items apply to CPU *replacement* scenarios. If you are replacing a blade server and *moving* existing CPUs to the new server, you do not have to separate the heatsink from the CPU. See [Additional CPU-Related Parts to Order with RMA Replacement Blade Server](#), on page 30.

- Scenario 1—You are reusing the existing heatsinks:
 - Heat sink cleaning kit (UCSX-HSCK=)
One cleaning kit can clean up to four CPUs.
 - Thermal interface material (TIM) kit for M5 servers (UCS-CPU-TIM=)
One TIM kit covers one CPU.
- Scenario 2—You are replacing the existing heatsinks:
(New heatsinks have a pre-applied pad of TIM.)
 - Heatsink for CPU 1: UCSB-HS-M5-F=
 - Heatsink for CPU 2: UCSB-HS-M5-R=
 - Heatsink cleaning kit (UCSX-HSCK=)
One cleaning kit can clean up to four CPUs.
- Scenario 3—You have a damaged CPU carrier (the plastic frame around the CPU):
 - CPU Carrier: UCS-M5-CPU-CAR=
 - #1 flat-head screwdriver (for separating the CPU from the heatsink)
 - Heatsink cleaning kit (UCSX-HSCK=)
One cleaning kit can clean up to four CPUs.
 - Thermal interface material (TIM) kit for M5 servers (UCS-CPU-TIM=)
One TIM kit covers one CPU.

A CPU heatsink cleaning kit is good for up to four CPU and heatsink cleanings. The cleaning kit contains two bottles of solution, one to clean the CPU and heatsink of old TIM and the other to prepare the surface of the heatsink.

New heatsink spares come with a pre-applied pad of TIM. It is important to clean any old TIM off of the CPU surface prior to installing the heatsinks. Therefore, even when you are ordering new heatsinks, you must order the heatsink cleaning kit.

Additional CPU-Related Parts to Order with RMA Replacement Blade Server

When a return material authorization (RMA) of the blade server is done, you move existing CPUs to the new server.



Note Unlike previous generation CPUs, the M5 server CPUs do not require you to separate the heatsink from the CPU when you *move* the CPU-heatsink assembly. Therefore, no additional heatsink cleaning kit or thermal-interface material items are required.

- The only tool required for moving a CPU/heatsink assembly is a T-30 Torx driver.

To move a CPU to a new blade server, use the procedure in [Moving an M5 Generation CPU, on page 30](#).

Moving an M5 Generation CPU

Tool required for this procedure: T-30 Torx driver



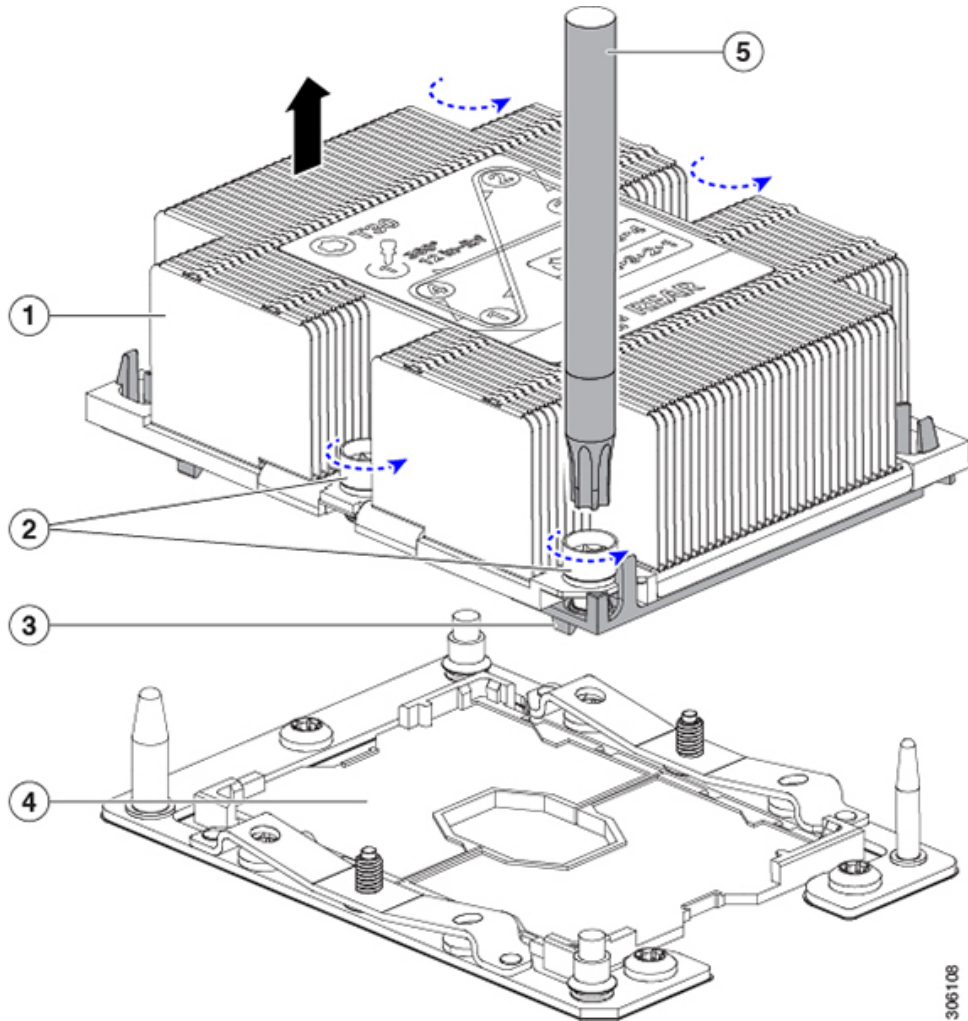
Caution When you receive a replacement server for an RMA, it includes dust covers on all CPU sockets. These covers protect the socket pins from damage during shipping. You must transfer these covers to the system that you are returning, as described in this procedure.

Procedure

- Step 1** When moving an M5 CPU to a new blade server, you do not have to separate the heatsink from the CPU. Perform the following steps:
- Use a T-30 Torx driver to loosen the four captive nuts that secure the assembly to the board standoffs.

Note Alternate loosening the heatsink nuts evenly so that the heatsink remains level as it is raised. Loosen the heatsink nuts in the order shown on the heatsink label: 4, 3, 2, 1.
 - Lift straight up on the CPU/heatsink assembly to remove it from the board.
 - Set the CPUs with heatsinks aside on an anti-static surface.

Figure 17: Removing the CPU/Heatsink Assembly



1	Heatsink	4	CPU socket on motherboard
2	Heatsink captive nuts (two on each side)	5	T-30 Torx driver
3	CPU carrier (below heatsink in this view)	-	

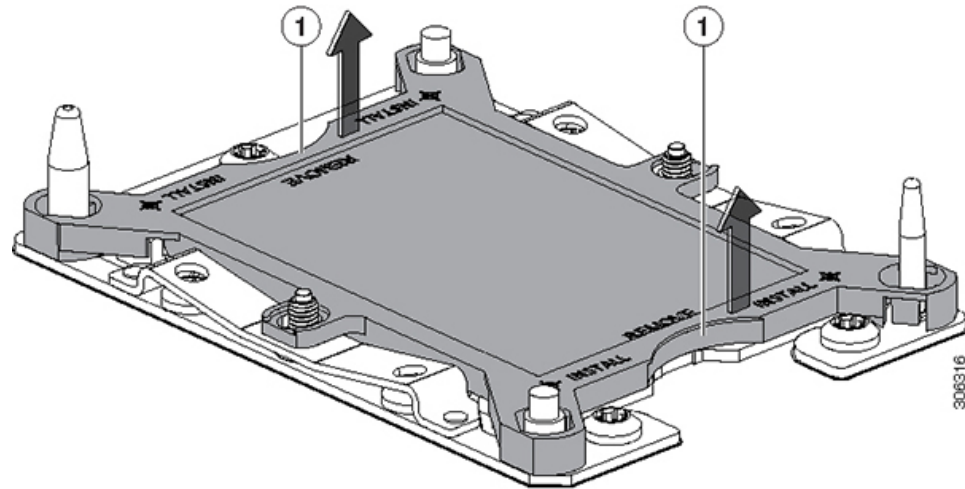
Step 2

Transfer the CPU socket covers from the new system to the system that you are returning:

- a) Remove the socket covers from the replacement system. Grasp the two recessed finger-grip areas marked "REMOVE" and lift straight up.

Note Keep a firm grasp on the finger-grip areas at both ends of the cover. Do not make contact with the CPU socket pins.

Figure 18: Removing a CPU Socket Dust Cover



1	Finger-grip areas marked "REMOVE" -	
---	-------------------------------------	--

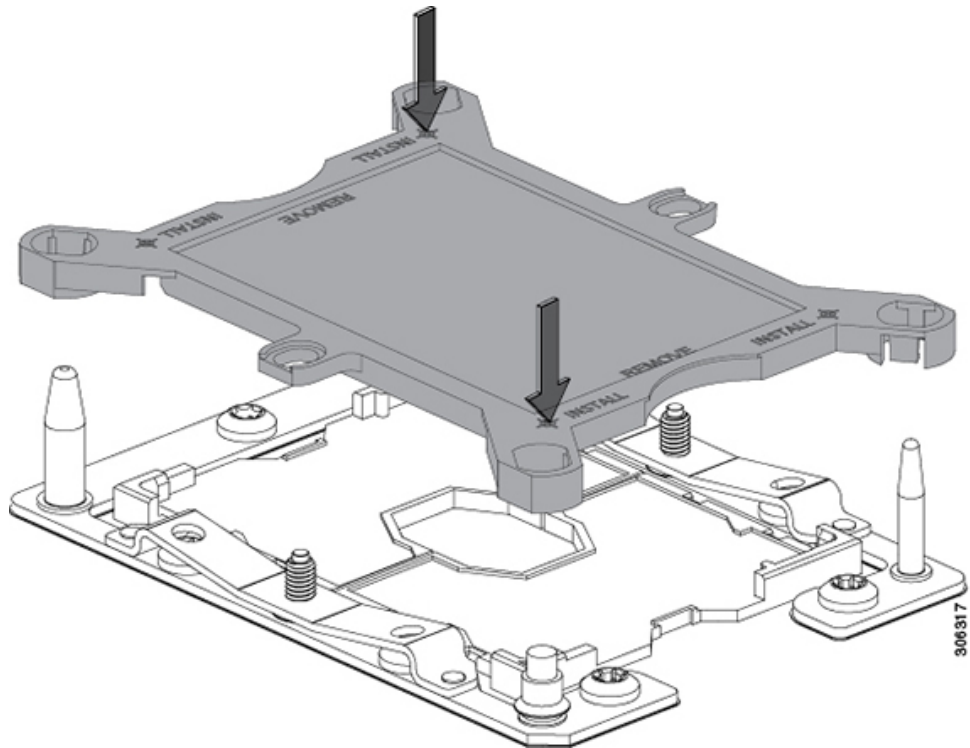
- b) With the wording on the dust cover facing up, set it in place over the CPU socket in the system that you are returning. Make sure that all alignment posts on the socket plate align with the cutouts on the cover.

Caution In the next step, do not press down anywhere on the dust cover except the two points described. Pressing elsewhere might damage the socket pins.

- c) Press down on the two circular markings next to the word "INSTALL" that are closest to the two threaded posts (see the following figure). Press until you feel and hear a click.

Note You must press until you feel and hear a click to ensure that the dust covers do not come loose during shipping.

Figure 19: Installing a CPU Socket Dust Cover



-	Press down on the two circular marks next to the word INSTALL.	-	
---	--	---	--

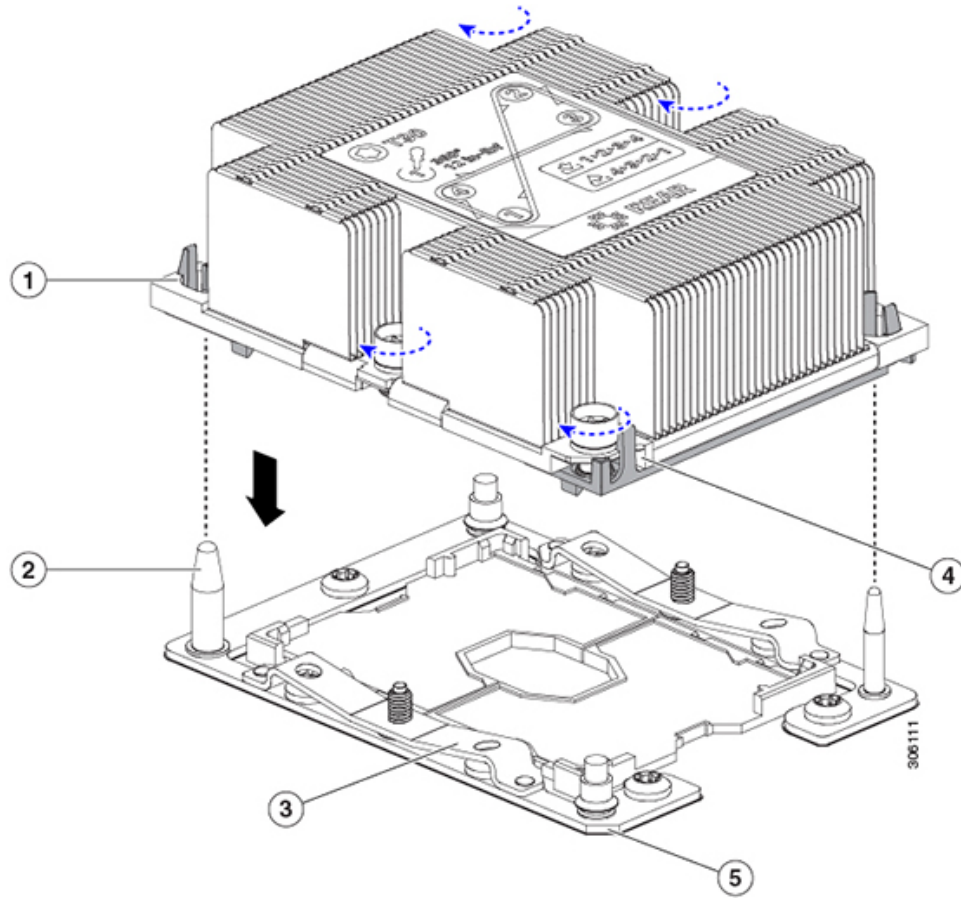
Step 3

Install the CPUs to the new system:

- a) On the new board, align the assembly over the CPU socket, as shown below.

Note the alignment features. The pin 1 angled corner on the heatsink must align with the pin 1 angled corner on the CPU socket. The CPU-socket posts must align with the guide-holes in the assembly.

Figure 20: Installing the Heatsink/CPU Assembly to the CPU Socket



1	Guide hole in assembly (two)	4	Angled corner on heatsink (pin 1 alignment feature)
2	CPU socket alignment post (two)	5	Angled corner on socket (pin 1 alignment feature)
3	CPU socket leaf spring	-	

- b) On the new board, set the heatsink with CPU assembly down onto the CPU socket.
- c) Use a T-30 Torx driver to tighten the four captive nuts that secure the heatsink to the board standoffs.

Note Alternate tightening the heatsink nuts evenly so that the heatsink remains level while it is lowered. Tighten the heatsink nuts in the order shown on the heatsink label: 1, 2, 3, 4. The captive nuts must be fully tightened so that the leaf springs on the CPU socket lie flat.

Replacing Memory DIMMs

The DIMMs that this blade server supports are updated frequently. A list of supported and available DIMMs is in Cisco UCS B480 M5 Specification Sheet.

Do not use any DIMMs other than those listed in the specification sheet. Doing so may irreparably damage the server and result in down time.

Memory Population Guidelines

This blade server contains 48 DIMM slots—12 per CPU.

**Caution**

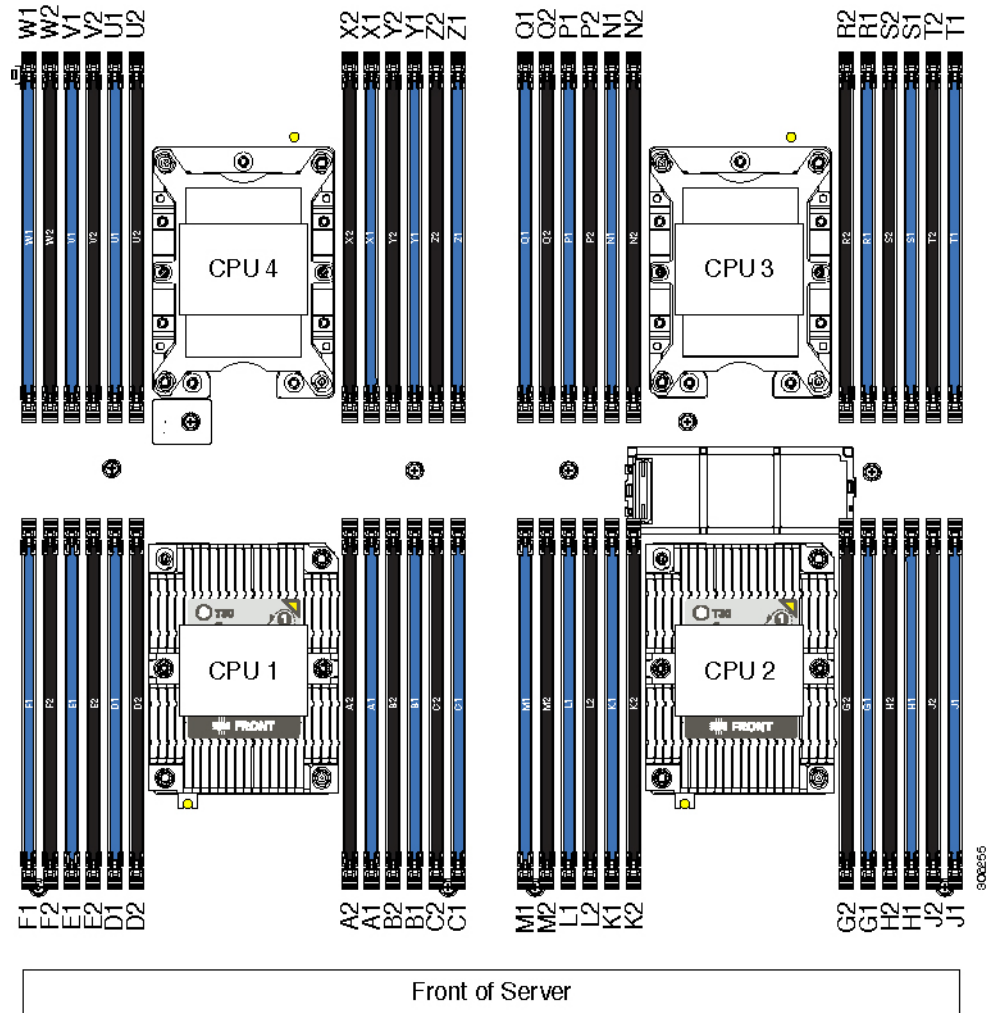
Only Cisco memory is supported. Third-party DIMMs are not tested or supported.

**Note**

When using 256 GB DDR DIMMs (UCS-ML-256G8RT-H) in this server, the blade-level power capping must be set to 1300 W. For information about blade-level power capping, see the *Power Capping and Power Management* chapter in the *Cisco UCS Manager Server Management Guide* for your release: [Cisco UCS Manager Configuration Guides](#).

- Each set of 12 DIMMs is arranged into six channels, where each channel has two DIMMs. Each channel is identified by a letter:
 - Channels A, B, C, D, E, and F are for CPU 1.
 - Channels G, H, J, K, L, and M are for CPU 2.
 - Channels N, P, Q, R, S, and T are for CPU 3.
 - Channels U, V, W, X, Y, and Z are for CPU 4.
- The maximum combined memory allowed in the 12 DIMM slots is 768 GB. To populate the 12 DIMM slots with more than 768 GB of combined memory, you need to use a CPU with a SKU that ends with an "M", for example, UCS-CPU-6134M.
- Do not mix DIMM types (LRDIMM, RDIMM,TSV-RDIMMs).
- The following figure shows how DIMMs and channels are physically laid out and numbered.

Figure 21: Physical Location of DIMM Slots



- A DIMM channel has either one or two DIMMs. For those channels with one DIMM, a DIMM blank must be installed. A slot cannot be empty. For installation instructions, see [Installing a DIMM or DIMM Blank](#), on page 37.
- For optimal performance, populate DIMMs in the order shown in the following table, depending on the number of CPUs and the number of DIMMs per CPU. If your server has four CPUs, balance DIMMs evenly across the four CPUs as shown in the table.



Note The table below lists recommended configurations. Using 5, 7, 9, 10, or 11 DIMMs per CPU is not recommended.

- The following table provides the DIMM population order:

Table 4: DIMM Population Order

Number of DIMMs per CPU (Recommended Configuration)	Populate CPU 1 Slot		Populate CPU 2 Slots		Populate CPU 3 Slots		Populate CPU 4 Slots	
	Blue #1 Slots	Black #2 Slots	Blue #1 Slots	Black #2 Slots	Blue #1 Slots	Black #2 Slots	Blue #1 Slots	Black #2 Slots
1	(A1)	-	(G1)	-	(N1)	-	(U1)	-
2	(A1, B1)	-	(G1, H1)	-	(N1, P1)	-	(U1, V1)	-
3	(A1, B1, C1)	-	(G1, H1, J1)	-	(N1, P1, Q1)	-	(U1, V1, W1)	-
4	(A1, B1); (D1, E1)	-	(G1, H1); (K1, L1)	-	(N1, P1); (R1, S1)	-	(U1, V1); (X1, Y1)	-
6	(A1, B1); (C1, D1); (E1, F1)	-	(G1, H1); (J1, K1); (L1, M1)	-	(N1, P1); (Q1, R1); (S1, T1)	-	(U1, V1); (W1, X1); (Y1, Z1)	-
8	(A1, B1); (D1, E1)	(A2, B2); (D2, E2)	(G1, H1); (K1, L1)	(G2, H2); (K2, L2)	(N1, P1); (R1, S1)	(N2, P2); (R2, S2)	(U1, V1); (X1, Y1)	(U2, V2); (X2, Y2)
12	(A1, B1); (C1, D1); (E1, F1)	(A2, B2); (C2, D2); (E2, F2)	(G1, H1); (J1, K1); (L1, M1)	(G2, H2); (J2, K2); (L2, M2)	(N1, P1); (Q1, R1); (S1, T1)	(N2, P2); (Q2, R2); (S2, T2)	(U1, V1); (W1, X1); (Y1, Z1)	(U2, V2); (W2, X2); (Y2, Z2)

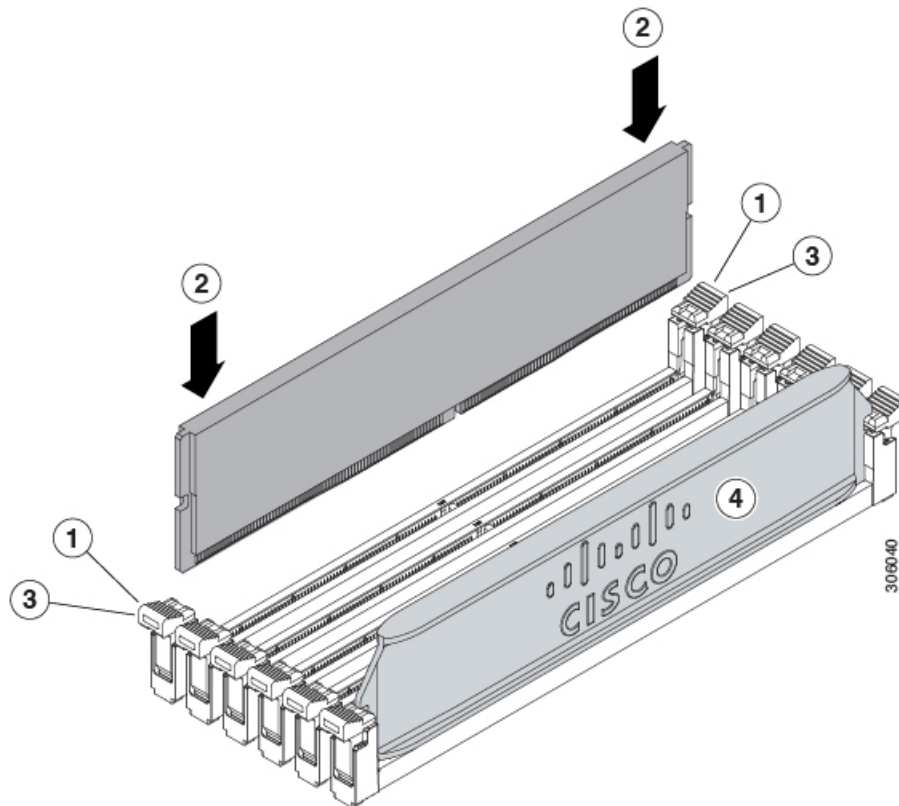
Installing a DIMM or DIMM Blank

To install a DIMM or a DIMM blank (UCS-DIMM-BLK=) into a slot on the blade server, follow these steps:

Procedure

-
- Step 1** Open both DIMM connector latches.
- Step 2** Press evenly on both ends of the DIMM until it clicks into place in its slot.
- Note** Ensure that the notch in the DIMM aligns with the slot. If the notch is misaligned, it is possible to damage the DIMM, the slot, or both.
- Step 3** Press the DIMM connector latches inward slightly to seat them fully.
- Step 4** Populate all slots with a DIMM or DIMM blank. A slot cannot be empty.

Figure 22: Installing Memory



Memory Performance

When considering the memory configuration of the blade server, there are several things to consider. For example:

- When mixing DIMMs of different densities (capacities), the highest density DIMM goes in slot 1 then in descending density.
- Besides DIMM population and choice, the selected CPU(s) can have some effect on performance.

Memory Mirroring and RAS

The Intel CPUs within the blade server support memory mirroring only when an even number of **channels** are populated with DIMMs. Furthermore, if memory mirroring is used, DRAM size is reduced by 50 percent for reasons of reliability.

Replacing Intel Optane DC Persistent Memory Modules

This topic contains information for replacing Intel Optane Data Center Persistent Memory modules (DCPMMs), including population rules. DCPMMs have the same form-factor as DDR4 DIMMs and they install to DIMM slots.



Note Intel Optane DC persistent memory modules require Second Generation Intel Xeon Scalable processors. You must upgrade the server firmware and BIOS to version 4.0(4) or later and install the supported Second Generation Intel Xeon Scalable processors before installing DCPMMs.



Caution DCPMMs and their sockets are fragile and must be handled with care to avoid damage during installation.



Note To ensure the best server performance, it is important that you are familiar with memory performance guidelines and population rules before you install or replace DCPMMs.

DCPMMs can be configured to operate in one of three modes:

- **Memory Mode (default):** The module operates as 100% memory module. Data is volatile and DRAM acts as a cache for DCPMMs. This is the factory default mode.
- **App Direct Mode:** The module operates as a solid-state disk storage device. Data is saved and is non-volatile.
- **Mixed Mode (25% Memory Mode + 75% App Direct):** The module operates with 25% capacity used as volatile memory and 75% capacity used as non-volatile storage.

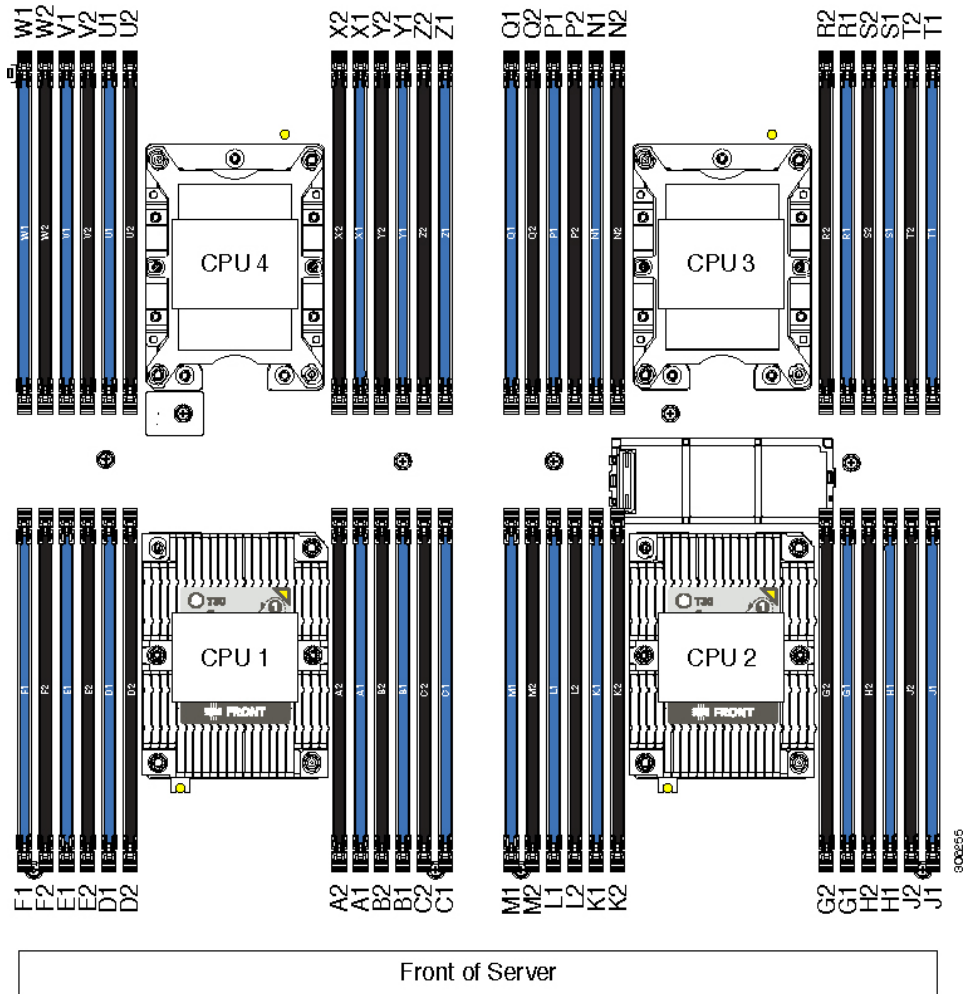
Intel Optane DC Persistent Memory Module Population Rules and Performance Guidelines

This topic describes the rules and guidelines for maximum memory performance when using Intel Optane DC persistent memory modules (DCPMMs) with DDR4 DIMMs.

DIMM Slot Numbering

The following figure shows the physical placement and numbering of the DIMM slots on the server board.

Figure 23: DIMM Slot Numbering



Configuration Rules

Observe the following rules and guidelines:

- To use DCPMMs in this server, four CPUs must be installed.
- Intel Optane DC persistent memory modules require Second Generation Intel Xeon Scalable processors. You must upgrade the server firmware and BIOS to version 4.0(4) or later and then install the supported Second Generation Intel Xeon Scalable processors before installing DCPMMs.
- When using DCPMMs in a server:
 - The DDR4 DIMMs installed in the server must all be the same size.
 - The DCPMMs installed in the server must all be the same size and must have the same SKU.
- The DCPMMs run at 2666 MHz. If you have 2933 MHz RDIMMs or LRDIMMs in the server and you add DCPMMs, the main memory speed clocks down to 2666 MHz to match the speed of the DCPMMs.
- Each DCPMM draws 18 W sustained, with a 20 W peak.

- The following table shows supported DCPMM configurations for this server. For dual CPU systems, fill the DIMM slots for CPU 1 and CPU 2 as shown, depending on which DCPMM:DRAM ratio you want to populate. For four-CPU systems, also fill the DIMM slots for CPU 3 and CPU 4 as shown.

Figure 24: Supported DCPMM Configurations for Quad-CPU Configurations

DIMM to DCPMM Count	CPU 1 (lower server node)											
	IMC1						IMC0					
	Channel 2		Channel 1		Channel 0		Channel 2		Channel 1		Channel 0	
	F2	F1	E2	E1	D2	D1	C2	C1	B2	B1	A2	A1
6 to 2		DIMM		DIMM	DCPMM	DIMM		DIMM		DIMM	DCPMM	DIMM
6 to 4		DIMM	DCPMM	DIMM	DCPMM	DIMM		DIMM	DCPMM	DIMM	DCPMM	DIMM
6 to 6	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM

DIMM to DCPMM Count	CPU 2 (lower server node)											
	IMC1						IMC0					
	Channel 2		Channel 1		Channel 0		Channel 2		Channel 1		Channel 0	
	M2	M1	L2	L1	K2	K1	J2	J1	H2	H1	G2	G1
6 to 2		DIMM		DIMM	DCPMM	DIMM		DIMM		DIMM	DCPMM	DIMM
6 to 4		DIMM	DCPMM	DIMM	DCPMM	DIMM		DIMM	DCPMM	DIMM	DCPMM	DIMM
6 to 6	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM

DIMM to DCPMM Count	CPU 3 (upper server node)											
	IMC1						IMC0					
	Channel 2		Channel 1		Channel 0		Channel 2		Channel 1		Channel 0	
	T2	T1	S2	S1	R2	R1	Q2	Q1	P2	P1	N2	N1
6 to 2		DIMM		DIMM	DCPMM	DIMM		DIMM		DIMM	DCPMM	DIMM
6 to 4		DIMM	DCPMM	DIMM	DCPMM	DIMM		DIMM	DCPMM	DIMM	DCPMM	DIMM
6 to 6	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM

DIMM to DCPMM Count	CPU 4 (upper server node)											
	IMC1						IMC0					
	Channel 2		Channel 1		Channel 0		Channel 2		Channel 1		Channel 0	
	Z2	Z1	Y2	Y1	X2	X1	W2	W1	V2	V1	U2	U1
6 to 2		DIMM		DIMM	DCPMM	DIMM		DIMM		DIMM	DCPMM	DIMM
6 to 4		DIMM	DCPMM	DIMM	DCPMM	DIMM		DIMM	DCPMM	DIMM	DCPMM	DIMM
6 to 6	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM	DCPMM	DIMM

3 073 80

Installing Intel Optane DC Persistent Memory Modules



Note DCPMM configuration is always applied to all DCPMMs in a region, including a replacement DCPMM. You cannot provision a specific replacement DCPMM on a preconfigured server.

Understand which mode your DCPMM is operating in. App Direct mode has some additional considerations in this procedure.



Caution Replacing a DCPMM in App-Direct mode requires all data to be wiped from the DCPMM. Make sure to backup or offload data before attempting this procedure.

Procedure

Step 1 For App Direct mode, backup the existing data stored in all Optane DIMMs to some other storage.

Step 2 For App Direct mode, remove the Persistent Memory policy which will remove goals and namespaces automatically from all Optane DIMMs.

Step 3 Remove an existing DCPMM:

- a) Decommission and power off the server.
- b) Remove the top cover from the server as described in [Removing a Blade Server Cover, on page 14](#).
- c) Slide the server out the front of the chassis.

Caution If you are moving DCPMMs with active data (persistent memory) from one server to another as in an RMA situation, each DCPMM must be installed to the identical position in the new server. Note the positions of each DCPMM or temporarily label them when removing them from the old server.

- d) Locate the DCPMM that you are removing, and then open the ejector levers at each end of its DIMM slot.

Step 4 Install a new DCPMM:

Note Before installing DCPMMs, see the population rules for this server: [Intel Optane DC Persistent Memory Module Population Rules and Performance Guidelines, on page 39](#).

- a) Align the new DCPMM with the empty slot on the motherboard. Use the alignment feature in the DIMM slot to correctly orient the DCPMM.
- b) Push down evenly on the top corners of the DCPMM until it is fully seated and the ejector levers on both ends lock into place.
- c) Replace the top cover to the server.
- d) Replace the server in the chassis.
- e) Wait for Cisco UCS Manager to complete its automatic discovery of the server.

Step 5 Perform post-installation actions:

Note If your Persistent Memory policy is Host Controlled, you must perform the following actions from the OS side.

- If the existing configuration is in 100% Memory mode, and the new DCPMM is also in 100% Memory mode (the factory default), the only action is to ensure that all DCPMMs are at the latest, matching firmware level.
- If the existing configuration is fully or partly in App-Direct mode and new DCPMM is also in App-Direct mode, then ensure that all DCPMMs are at the latest matching firmware level and also re-provision the DCPMMs by creating a new goal.
 - For App Direct mode, reapply the Persistent Memory policy.
 - For App Direct mode, restore all the offloaded data to the DCPMMs.

- If the existing configuration and the new DCPMM are in different modes, then ensure that all DCPMMs are at the latest matching firmware level and also re-provision the DCPMMs by creating a new goal.

There are a number of tools for configuring goals, regions, and namespaces.

- To use the server's BIOS Setup Utility, see [Server BIOS Setup Utility Menu for DCPMM](#), on page 43.
- To use Cisco IMC or Cisco UCS Manager, see the [Cisco UCS: Configuring and Managing Intel Optane DC Persistent Memory Modules](#) guide.

Server BIOS Setup Utility Menu for DCPMM



Caution Potential data loss: If you change the mode of a currently installed DCPMM from App Direct or Mixed Mode to Memory Mode, any data in persistent memory is deleted.

DCPMMs can be configured by using the server's BIOS Setup Utility, Cisco IMC, Cisco UCS Manager, or OS-related utilities.

- To use the BIOS Setup Utility, see the section below.
- To use Cisco IMC, see the configuration guides for Cisco IMC 4.0(4) or later: [Cisco IMC CLI and GUI Configuration Guides](#)
- To use Cisco UCS Manager, see the configuration guides for Cisco UCS Manager 4.0(4) or later: [Cisco UCS Manager CLI and GUI Configuration Guides](#)

The server BIOS Setup Utility includes menus for DCPMMs. They can be used to view or configure DCPMM regions, goals, and namespaces, and to update DCPMM firmware.

To open the BIOS Setup Utility, press **F2** when prompted during a system boot.

The DCPMM menu is on the Advanced tab of the utility:

Advanced > Intel Optane DC Persistent Memory Configuration

From this tab, you can access other menu items:

- **DIMMs**: Displays the installed DCPMMs. From this page, you can update DCPMM firmware and configure other DCPMM parameters.
 - Monitor health
 - Update firmware
 - Configure security
 - You can enable security mode and set a password so that the DCPMM configuration is locked. When you set a password, it applies to all installed DCPMMs. Security mode is disabled by default.
 - Configure data policy
- **Regions**: Displays regions and their persistent memory types. When using App Direct mode with interleaving, the number of regions is equal to the number of CPU sockets in the server. When using

App Direct mode without interleaving, the number of regions is equal to the number of DCPMMs in the server.

From the Regions page, you can configure memory goals that tell the DCPMM how to allocate resources.

- `Create goal config`

- **Namespaces:** Displays namespaces and allows you to create or delete them when persistent memory is used. Namespaces can also be created when creating goals. A namespace provisioning of persistent memory applies only to the selected region.

Existing namespace attributes such as the size cannot be modified. You can only add or delete namespaces.

- **Total capacity:** Displays the total resource allocation across the server.

Updating the DCPMM Firmware Using the BIOS Setup Utility

You can update the DCPMM firmware from the BIOS Setup Utility if you know the path to the .bin files. The firmware update is applied to all installed DCPMMs.

1. Navigate to **Advanced > Intel Optane DC Persistent Memory Configuration > DIMMs > Update firmware**
2. Under **File:**, provide the file path to the .bin file.
3. Select **Update**.

Installing a Virtual Interface Card in the mLOM Slot

Cisco VIC cards supported in the mLOM slot:

Cisco VIC cards supported in the mLOM slot (use the procedure below):

- Cisco UCS VIC 1340
- Cisco UCS VIC 1440

Cisco VIC cards supported in the optional VIC port expander (use the procedure in [Installing a Rear Mezzanine Module in Addition to the mLOM VIC, on page 45](#)):

- Cisco UCS VIC 1380
- Cisco UCS VIC 1480

To install Cisco VIC card in the blade server, follow these steps.

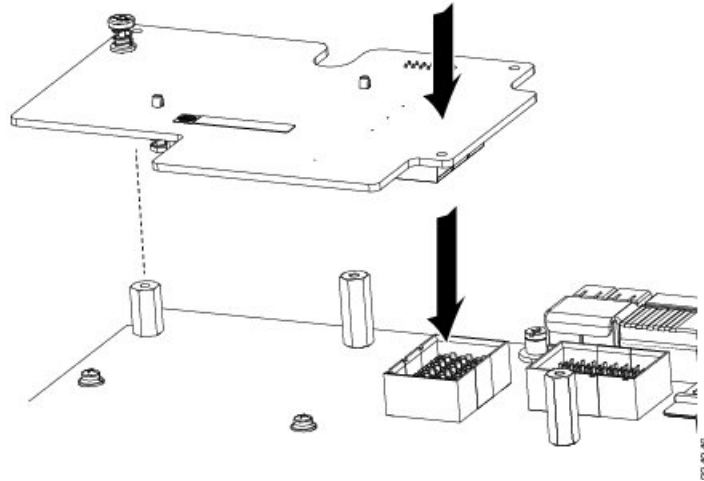
Procedure

-
- Step 1** If a rear mezzanine module is installed, remove it to provide access to the mLOM slot.
 - Step 2** Position the VIC connector above the motherboard connector and align the captive screw to the standoff post on the motherboard.
 - Step 3** Firmly press the VIC connector into the motherboard connector where “PRESS HERE TO INSTALL” is stated.

Step 4 Tighten the captive screw.

Tip To remove a VIC, reverse the above procedure. You might find it helpful when removing the connector from the motherboard to gently rock the board along the length of the connector until it loosens.

Figure 25: Installing a VIC in the mLOM Slot



Installing a Rear Mezzanine Module in Addition to the mLOM VIC

All supported rear mezzanine modules have a common installation process. A list of currently supported and available rear mezzanine modules for this server is in the *Cisco UCS B200 M5 Blade Server Specification Sheet* on the [Cisco UCS B-Series Blade Servers Data Sheets](#) page. The rear mezzanine slots can be used for the VIC port expander, the NVIDIA P6 GPU, and non-I/O mezzanine cards.

Cisco VIC cards supported in the rear VIC port expander (use the procedure in this section):

- Cisco UCS VIC 1380
- Cisco UCS VIC 1480

Cisco VIC cards supported in the mLOM slot (use the procedure in [Installing a Virtual Interface Card in the mLOM Slot, on page 44](#)):

- Cisco UCS VIC 1340
- Cisco UCS VIC 1440

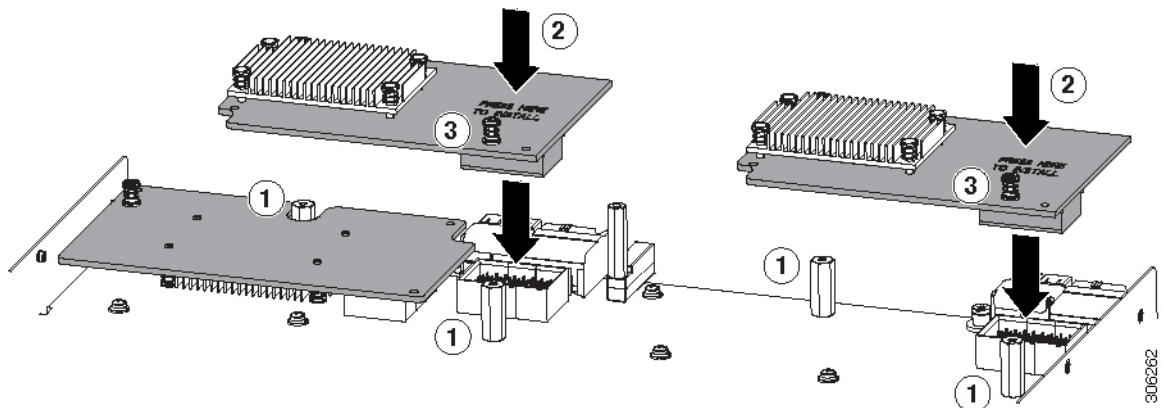


Note If you are switching from one type of rear mezzanine module to another, before you physically perform the switch, download the latest device drivers and load them into the server's operating system. For more information, see the firmware management chapter of one of the Cisco UCS Manager software configuration guides.

Procedure

- Step 1** Position the rear mezzanine module above the motherboard connector (callout 1) and align the two rear mezzanine module captive screws to the standoff posts on the motherboard.
- Step 2** Firmly press the rear mezzanine module into the motherboard connector (callout 2) where “PRESS HERE TO INSTALL” is stated.
- Step 3** Tighten the two rear mezzanine module captive screws (callout 3).
- Tip** Removing a rear mezzanine module is the reverse of installing it. You might find it helpful when removing the rear mezzanine module from the motherboard to gently rock the rear mezzanine module along the length of the motherboard connector until it loosens.

Figure 26: Installing a Rear Mezzanine Module



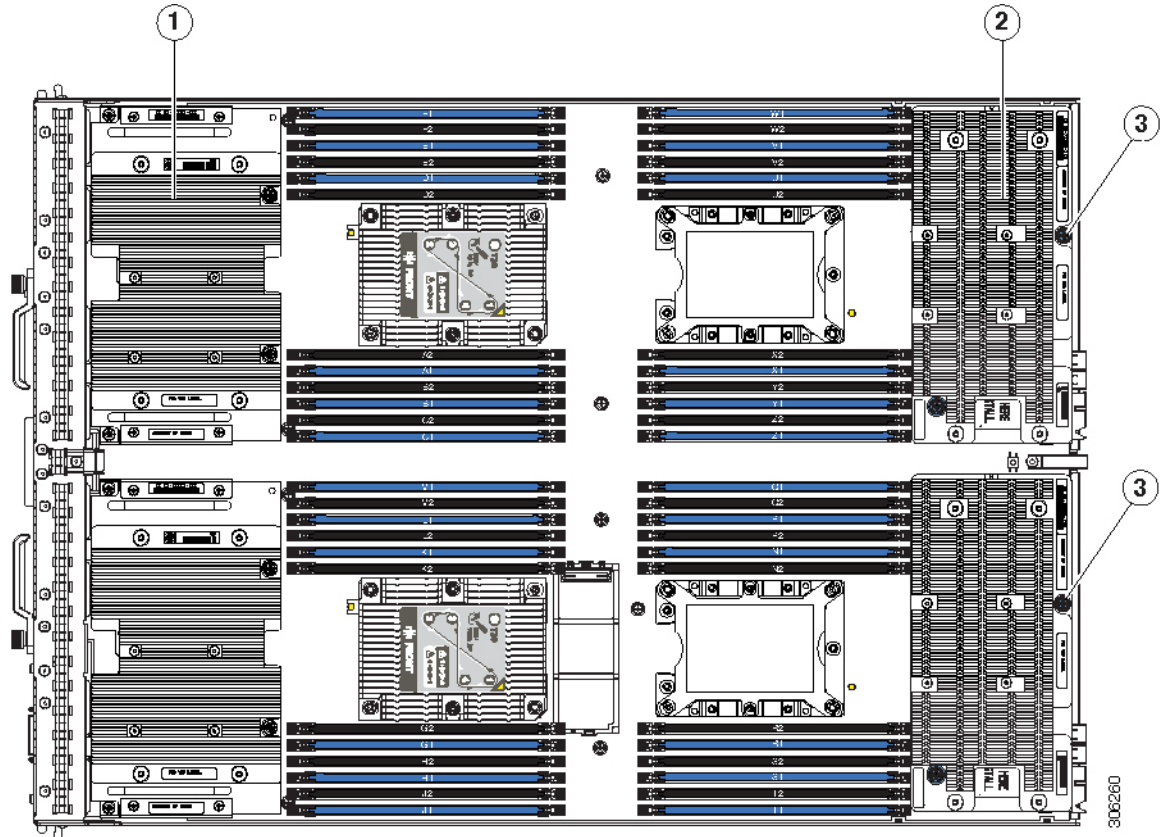
NVIDIA P6 Graphics Processing Unit

The NVIDIA P6 graphics processing unit (GPU) card provides graphics and computing capabilities to the server. There are two supported versions of the NVIDIA P6 GPU card:

- UCSB-GPU-P6-F can be installed only in the front mezzanine slot of the server. For installation instructions, see [Installing an NVIDIA GPU Card in the Front of the Server, on page 47](#).
- UCSB-GPU-P6-R can be installed only in the rear mezzanine slot of the server. For installation instructions, see [Installing an NVIDIA GPU Card in the Rear of the Server, on page 51](#).

The following figure shows the installed NVIDIA P6 GPU in the front and rear mezzanine slots.

Figure 27: NVIDIA GPU Installed in the Front and Rear Mezzanine Slots

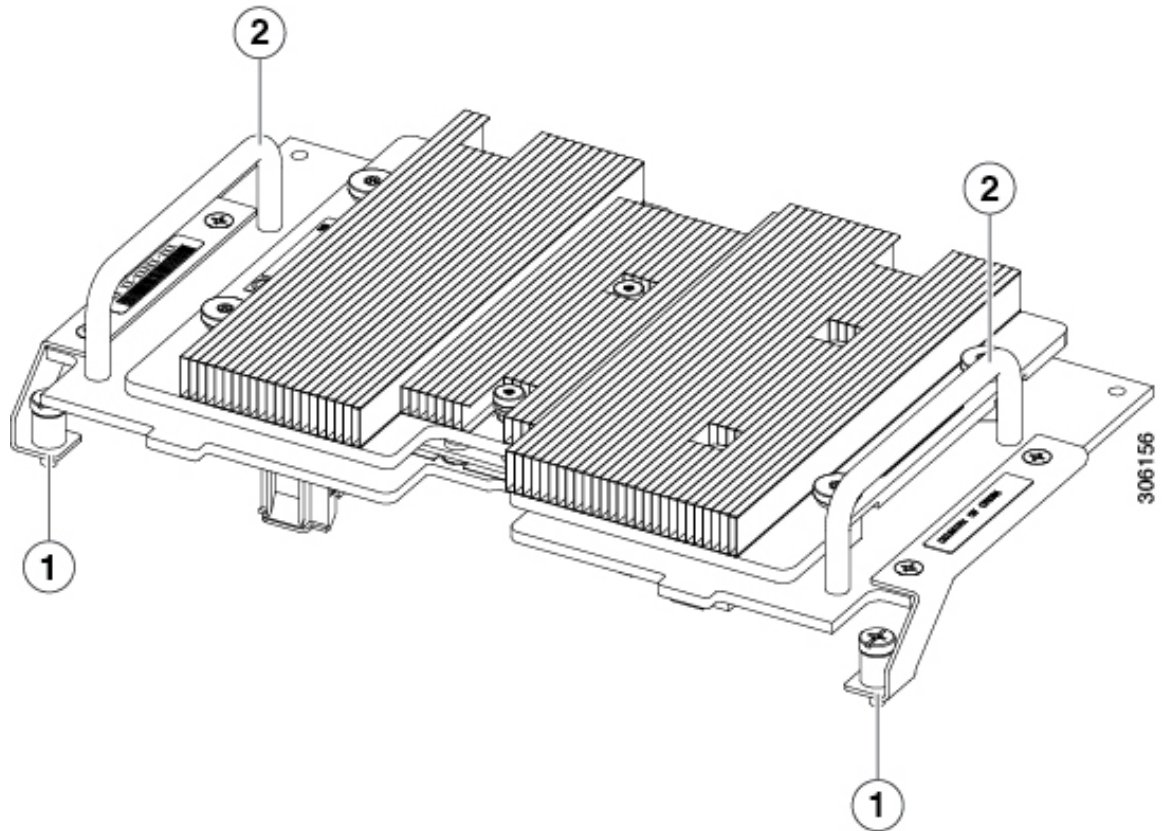


1	Front GPU	2	Rear GPU
3	Custom standoff screw	-	

Installing an NVIDIA GPU Card in the Front of the Server

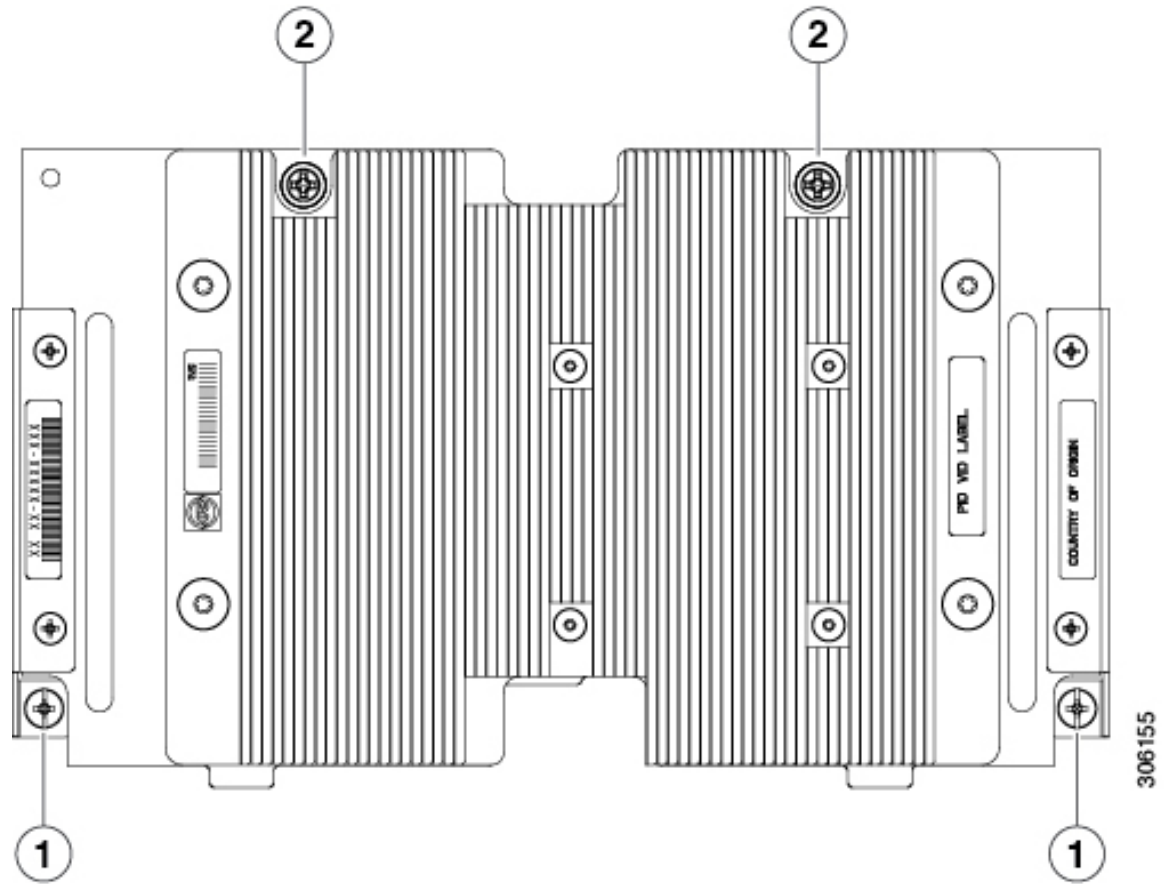
The following figure shows the front NVIDIA P6 GPU (UCSB-GPU-P6-F).

Figure 28: NVIDIA P6 GPU That Installs in the Front of the Server



<p>1</p>	<p>Leg with thumb screw that attaches to the server motherboard at the front</p>	<p>2</p>	<p>Handle to press down on when installing the GPU</p>
-----------------	--	-----------------	--

Figure 29: Top Down View of the NVIDIA P6 GPU for the Front of the Server



1	Leg with thumb screw that attaches to the server motherboard	2	Thumb screw that attaches to a standoff below
---	--	---	---

To install the NVIDIA GPU, follow these steps:

Before you begin

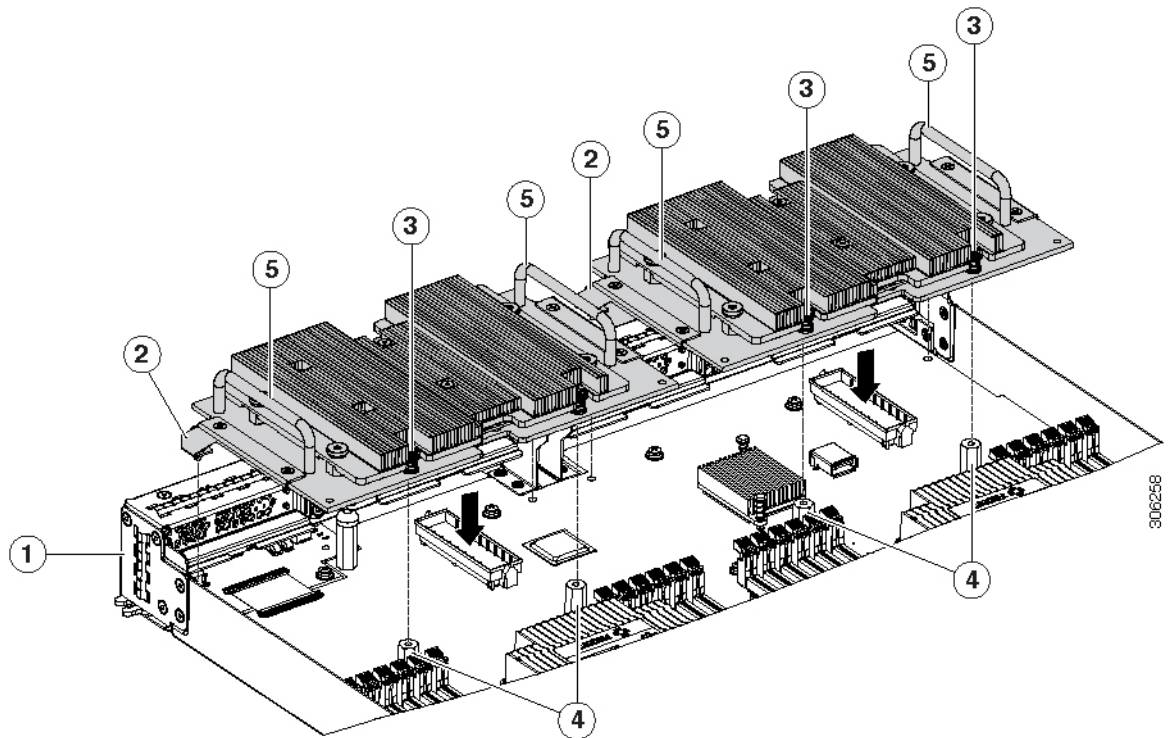
Before installing the NVIDIA P6 GPU (UCSB-GPU-P6-F) in the front mezzanine slot:

- Upgrade the Cisco UCS domain that the GPU will be installed into to a version of Cisco UCS Manager that supports this card. Refer to the latest version of the *Release Notes for Cisco UCS Software* at the following URL for information about supported hardware: <http://www.cisco.com/c/en/us/support/servers-unified-computing/ucs-manager/products-release-notes-list.html>.
- Remove the front mezzanine storage module if it is present. You cannot use the storage module in the front mezzanine slot when the NVIDIA P6 GPU is installed in the front of the server.

Procedure

- Step 1** Position the GPU in the correct orientation to the front of the server (callout 1) as shown in the following figure.
- Step 2** Install the GPU into the server. Press down on the handles (callout 5) to firmly secure the GPU.
- Step 3** Tighten the thumb screws (callout 3) at the back of the GPU with the standoffs (callout 4) on the motherboard.
- Step 4** Tighten the thumb screws on the legs (callout 2) to the motherboard.
- Step 5** Install the drive blanking panels.

Figure 30: Installing the NVIDIA GPU in the Front of the Server

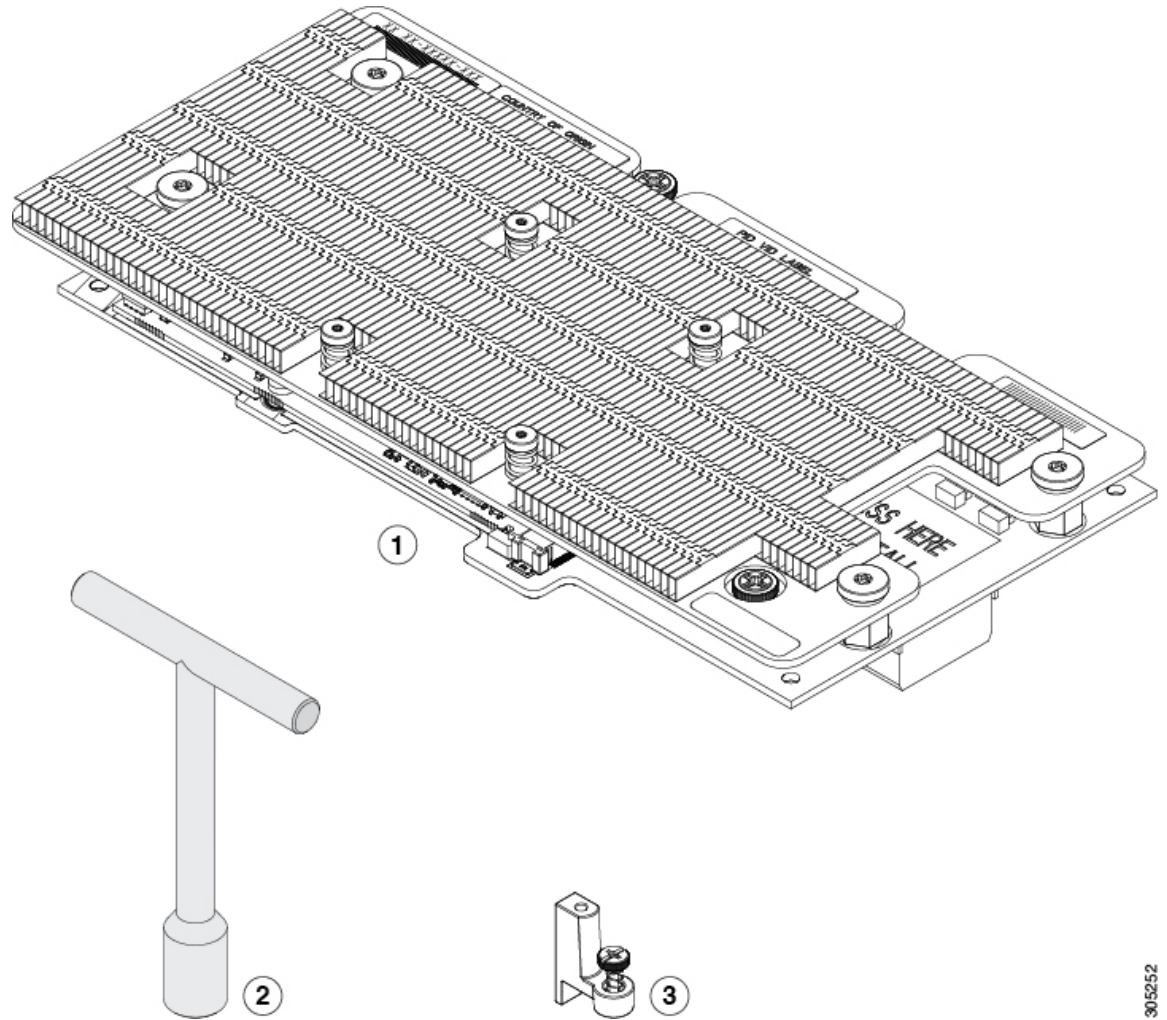


1	Front of the server	2	Leg with thumb screw that attaches to the motherboard
3	Thumbscrew to attach to standoff below	4	Standoff on the motherboard
5	Handle to press down on to firmly install the GPU	–	

Installing an NVIDIA GPU Card in the Rear of the Server

If you are installing the UCSB-GPU-P6-R to a server in the field, the option kit comes with the GPU itself (CPU and heatsink), a T-shaped installation wrench, and a custom standoff to support and attach the GPU to the motherboard. The following figure shows the three components of the option kit.

Figure 31: NVIDIA P6 GPU (UCSB-GPU-P6-R) Option Kit



1	NVIDIA P6 GPU (CPU and heatsink)	2	T-shaped wrench
3	Custom standoff	-	

Before you begin

Before installing the NVIDIA P6 GPU (UCSB-GPU-P6-R) in the rear mezzanine slot:

- Upgrade the Cisco UCS domain that the GPU will be installed into to a version of Cisco UCS Manager that supports this card. Refer to the latest version of the *Release Notes for Cisco UCS Software* at the

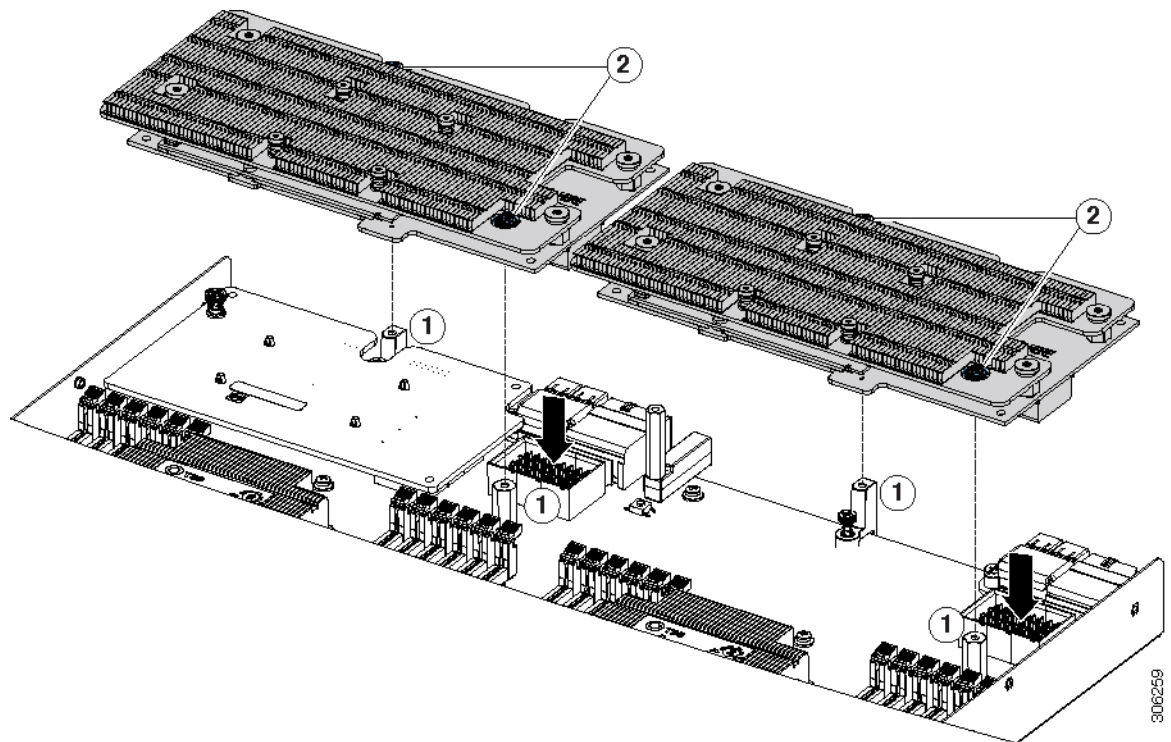
following URL for information about supported hardware: <http://www.cisco.com/c/en/us/support/servers-unified-computing/ucs-manager/products-release-notes-list.html>.

- Remove any other card, such as a VIC 1480, VIC 1380, or VIC port expander card from the rear mezzanine slot. You cannot use any other card in the rear mezzanine slot when the NVIDIA P6 GPU is installed.

Procedure

- Step 1** Use the T-shaped wrench that comes with the GPU to remove the existing standoff at the back end of the motherboard.
- Step 2** Install the custom standoff in the same location at the back end of the motherboard.
- Step 3** Position the CPU over the connector on the motherboard and align all the captive screws to the standoff posts (callout 1).
- Step 4** Tighten the captive screws (callout 2).

Figure 32: Installing the NVIDIA P6 GPU in the Rear Mezzanine Slot



Enabling the Trusted Platform Module

The Trusted Platform Module (TPM) is a component that can securely store artifacts used to authenticate the server. These artifacts can include passwords, certificates, or encryption keys. A TPM can also be used to store platform measurements that help ensure that the platform remains trustworthy. Authentication (ensuring

that the platform can prove that it is what it claims to be) and attestation (a process helping to prove that a platform is trustworthy and has not been breached) are necessary steps to ensure safer computing in all environments. It is a requirement for the Intel Trusted Execution Technology (TXT) security feature, which must be enabled in the BIOS settings for a server equipped with a TPM.

The Cisco UCS B200 M5 supports two options of TPM:

- The TPM 1.2.
- The TPM 2.0, UCSX-TPM2-002B(=), which is compliant with Federal Information Processing (FIPS) Standard 140-2. FIPS support has existed, but FIPS 140-2 is now supported.

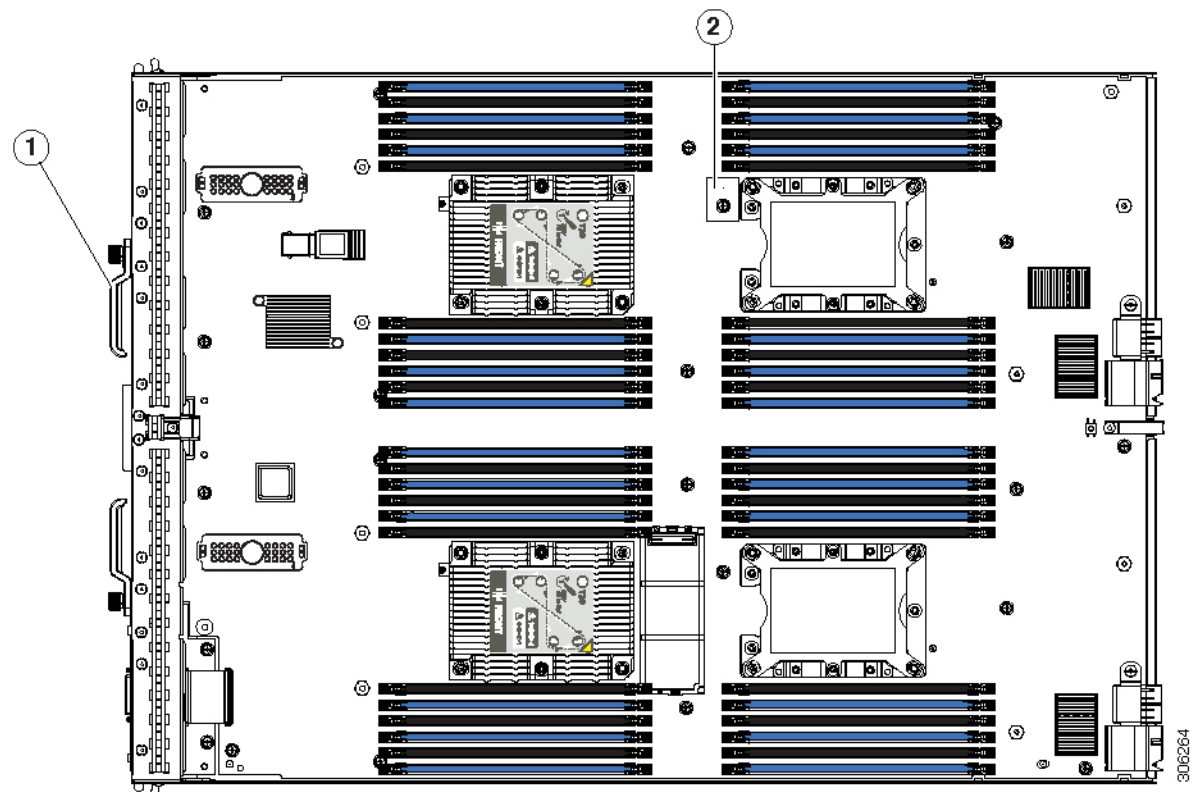
Procedure

Step 1

Install the TPM hardware.

- Decommission, power off, and remove the blade server from the chassis.
- Remove the top cover from the server as described in [Removing a Blade Server Cover, on page 14](#)
- Install the TPM to the TPM socket on the server motherboard and secure it using the one-way screw that is provided. See the figure below for the location of the TPM socket.
- Return the blade server to the chassis and allow it to be automatically reacknowledged, reassociated, and recommissioned.
- Continue with enabling TPM support in the server BIOS in the next step.

Figure 33: TPM Socket Location



1	Front of server	2	TPM socket on motherboard
----------	-----------------	----------	---------------------------

Step 2 Enable TPM Support in the BIOS.

- a) In the Cisco UCS Manager Navigation pane, click the **Servers** tab.
- b) On the Servers tab, expand **Servers > Policies**.
- c) Expand the node for the organization where you want to configure the TPM.
- d) Expand BIOS Policies and select the BIOS policy for which you want to configure the TPM.
- e) In the Work pane, click the **Advanced** tab.
- f) Click the **Trusted Platform** sub-tab.
- g) To enable TPM support, click **Enable** or **Platform Default**.
- h) Click **Save Changes**.
- i) Continue with the next step.

Step 3 Enable TXT Support in the BIOS Policy.

Follow the procedures in the [Cisco UCS Manager Configuration Guide](#) for the release that is running on the server.

Removing the Trusted Platform Module (TPM)

The TPM module is attached to the printed circuit board assembly (PCBA). You must disconnect the TPM module from the PCBA before recycling the PCBA. The TPM module is secured to a threaded standoff by a tamper-resistant screw. If you do not have the correct tool for the screw, you can use a pair of pliers to remove the screw.

Before you begin



Note **For Recyclers Only!** This procedure is not a standard field-service option. This procedure is for recyclers who will be reclaiming the electronics for proper disposal to comply with local eco design and e-waste regulations.

To remove the Trusted Platform Module (TPM), the following requirements must be met for the server:

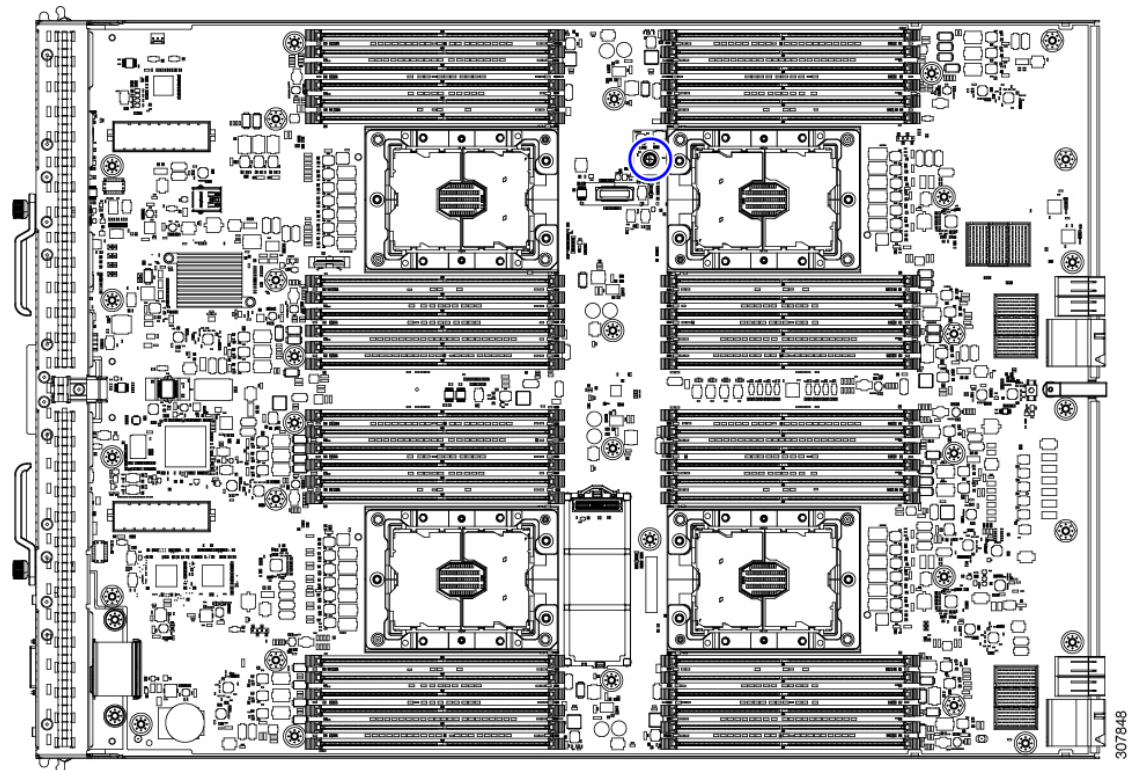
- It must be disconnected from facility power.
- It must be removed from the equipment rack.
- The top cover must be removed. If the top cover is not removed, see [Removing a Blade Server Cover, on page 14](#).

Procedure

Step 1 Locate the TPM module.

The following illustration shows the location of the TPM module's screw.

Figure 34: Screw Location for Removing the TPM Module



- Step 2** Using the pliers, grip the head of the screw and turn it counter clockwise until the screw releases.
- Step 3** Remove the TPM module and dispose of it properly.

What to do next

Remove and dispose of the PCB Assembly. See [Recycling the PCB Assembly \(PCBA\)](#), on page 71.

Mini Storage Module

The server has a mini-storage module option that plugs into a motherboard socket to provide additional internal storage. The mini-storage module can be one of the following types:

- An SD card module that supports up to two SD cards.
- An M.2 SSD module that supports up to two SATA M.2 SSDs.



Note The Cisco IMC firmware does not include an out-of-band management interface for the M.2 drives installed in the M.2 version of this mini-storage module (UCS-MSTOR-M2). The M.2 drives are not listed in Cisco IMC inventory, nor can they be managed by Cisco IMC. This is expected behavior.

Replacing the Mini-Storage Module Carrier

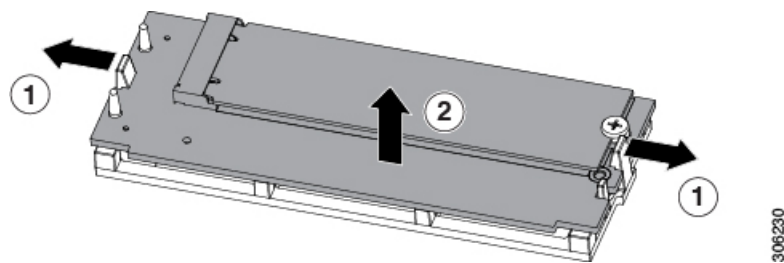
This topic describes how to remove and replace a mini-storage module carrier. The carrier has one media socket on its top and one socket on its underside. Use the following procedure for any type of mini-storage module carrier (SD card or M.2 SSD).

Procedure

Step 1 Remove the carrier from the server:

- a) Press out on the securing clips to disengage the module from the server board socket.
- b) Pull straight up on the storage module to remove it.

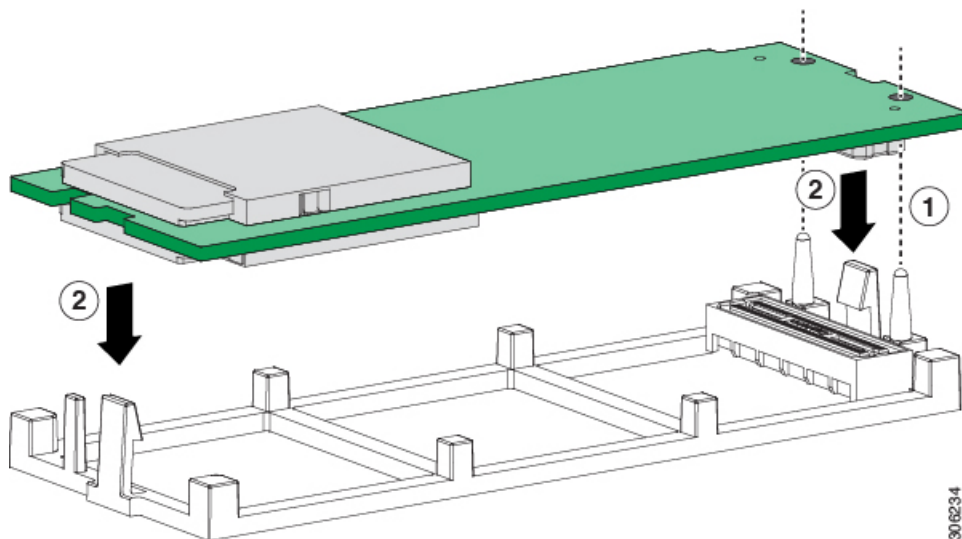
Figure 35: Removing the Mini-Storage Module Carrier



Step 2 Install a carrier to the server:

- a) Align the two holes on the carrier with the holder pins.
- b) Push the carrier into the holder on both ends, making sure the securing clips snap in. Push on the four corners of the carrier to fully seat it.

Figure 36: Installing the Mini-Storage Module Carrier



Replacing an M.2 SSD

This task describes how to remove an M.2 SSD from the mini-storage module carrier.

Population Rules For Mini-Storage M.2 SSDs

- You can use one or two M.2 SSDs in the carrier.
- M.2 socket 1 is on the top side of the carrier; M.2 socket 2 is on the underside of the carrier (the same side as the carrier's connector to the server board socket).
- Dual SATA M.2 SSDs can be configured in a RAID 1 array through the BIOS Setup Utility's embedded SATA RAID interface. See [Embedded SATA RAID Controller, on page 58](#).



Note The M.2 SSDs cannot be controlled by a hardware RAID controller card.

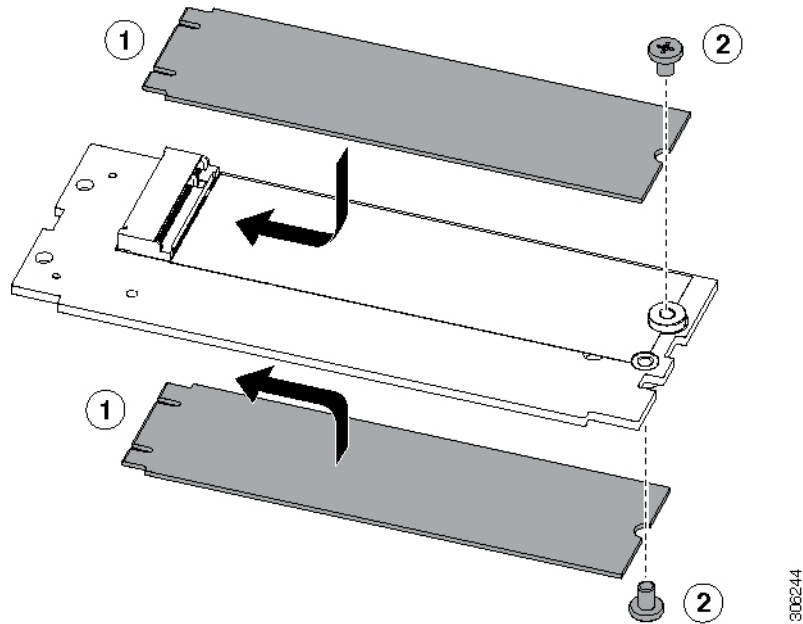


Note The embedded SATA RAID controller requires that the server is set to boot in UEFI mode rather than Legacy mode.

Procedure

- Step 1** Remove an M.2 SSD from the carrier:
- a) Using a #1 Phillips-head screwdriver, loosen the single screw that secures the M.2 SSD to the mini-storage module carrier.
 - b) Pull the M.2 SSD from the socket on the carrier.
- Step 2** Install an M.2 SSD:
- a) Align the gold fingers on the M.2 card with the module connector on the top of the M.2 mini-storage module, and then fully push the M.2 card into the module connector.
 - b) Using a #1 Phillips-head screwdriver, install the single screw to secure the M.2 card to the M.2 mini-storage module.

Figure 37: Installing M.2 SSDs to the Carrier



Embedded SATA RAID Controller

The server includes an embedded SATA MegaRAID controller that can be used to control internal SATA M.2 drives. This controller supports RAID levels 0 and 1.



Note The VMware ESX/ESXi operating system is not supported with the embedded SATA MegaRAID controller in SW RAID mode. You can use VMWare in AHCI mode.



Note The Microsoft Windows Server 2016 Hyper-V hypervisor is supported for use with the embedded MegaRAID controller in SW RAID mode, but all other hypervisors are not supported. All Hypervisors are supported in AHCI mode.



Note You cannot control the M.2 SATA SSDs in the server with a HW RAID controller.

Embedded SATA RAID Requirements

The embedded SATA RAID controller requires the following items:

- The embedded SATA RAID controller must be enabled in Cisco UCS Manager.

- M.2 mini-storage module with two SATA M.2 SSDs.
- The software RAID controller requires UEFI boot mode; legacy boot mode is not supported.
- (Optional) LSI MegaSR drivers for Windows or Linux.
- If you use an embedded RAID controller with Linux, both the pSATA and the sSATA controller must be set to LSI SW RAID mode.

Embedded SATA RAID Controller Considerations

Note the following considerations:

- The default setting for this embedded controller hub is SATA RAID 0 and 1 support for two M.2 SATA drives. The hub is divided into two SATA controllers that have different functions. See [Embedded SATA RAID: Two SATA Controllers, on page 59](#).
- When you order the server with this embedded controller, the controller is enabled. Instructions for enabling the controller are included for the case in which a server is reset to defaults. See [Enabling SATA Mode, on page 59](#).
- The required drivers for this controller are already installed and ready to use. However, if you will use this controller with Windows or Linux, you must download and install additional drivers for those operating systems. See [Installing LSI MegaSR Drivers For Windows and Linux, on page 60](#).

Embedded SATA RAID: Two SATA Controllers

The embedded RAID platform controller hub (PCH) is split into two controllers: primary SATA (pSATA) and secondary SATA (sSATA). These two controllers are seen as separate RAID controllers in Cisco UCS Manager:

- The primary pSATA controller is disabled.
- The secondary sSATA controller controls two internal M.2 SATA drives, when they are present in the M.2 mini-storage module option.
- Each controller is listed separately in Cisco UCS Manager. You can enable or disable the sSATA controller in Cisco UCS Manager. See [Enabling SATA Mode, on page 59](#).

Enabling SATA Mode

Perform this procedure in Cisco UCS Manager.

Procedure

- Step 1** Set the SATA mode.
- a) To change the M.2 state for the sSATA controller, change it in the storage sub-profile of the service profile that is assigned to the blade server. Choices are:
 - LSI SW RAID SWR—Enable the embedded sSATA RAID controller for control of internal SATA M.2 drives.

- AHCI—Enable control of the internal SATA M.2 drives by AHCI through your OS rather than the embedded RAID controller.
- Disabled—Disable the embedded sSATA RAID controller.

Step 2 Press **F10** to save your changes and exit.

Accessing the LSI Software RAID Configuration Utility

To configure RAID settings for the embedded SATA RAID controller, use the utility that is built into the BIOS. Each controller is controlled by its own instance of the utility.

Procedure

- Step 1** Boot the server and press **F2** when prompted to enter the BIOS Setup utility.
- Step 2** Choose the **Advanced** tab.
- Step 3** Select **LSI Software RAID Configuration Utility (sSATA)** for the sSATA controller.
-

Installing LSI MegaSR Drivers For Windows and Linux



Note The required drivers for this controller are already installed and ready to use. However, if you will use this controller with Windows or Linux, you must download and install additional drivers for those operating systems.

This section explains how to install the LSI MegaSR drivers for the following supported operating systems:

- Microsoft Windows Server
- Red Hat Enterprise Linux (RHEL)
- SUSE Linux Enterprise Server (SLES)

For the specific supported OS versions, see the [Hardware and Software Compatibility Matrix](#) for your server release.

Downloading the MegaSR Drivers

The MegaSR drivers are included in the B-Series driver ISO for your server and OS.

Procedure

- Step 1** Find the drivers ISO file download for your server online and download it to a temporary location on your workstation:
- See the following URL: <http://www.cisco.com/cisco/software/navigator.html>.

- b) Type the name of your server in the **Select a Product** search field and then press **Enter**.
- c) Click **Unified Computing System (UCS) Drivers**.
- d) Click the release number that you are downloading.
- e) Click the Download icon to download the drivers ISO file.

Step 2 Continue through the subsequent screens to accept the license agreement and then browse to the location where you want to save the driver ISO file.

Microsoft Windows Server Drivers

Installing Microsoft Windows Server Drivers

The Windows Server operating system automatically adds the driver to the registry and copies the driver to the appropriate directory.

Before you begin

Before you install this driver on the sSATA embedded controller, you must configure a RAID drive group.

To access the configuration utility, open the BIOS Setup Utility, go to the **Advanced** tab, and then choose the utility instance for the sSATA embedded controller: **LSI Software RAID Configuration Utility (sSATA)**.

Procedure

- Step 1** Download the Cisco UCS B-Series drivers' ISO, as described in [Downloading the MegaSR Drivers, on page 60](#).
- Step 2** Prepare the drivers on a USB thumb drive:
 - a) Burn the ISO image to a disk.
 - b) Browse the contents of the drivers folders to the location of the embedded MegaRAID drivers:
/ <OS>/Storage/Intel/B600/
 - c) Expand the Zip file, which contains the folder with the MegaSR driver files.
 - d) Copy the expanded folder to a USB thumb drive.
- Step 3** Start the Windows driver installation using one of the following methods:
 - To install from local media, connect an external USB DVD drive to the server (if the server does not have a DVD drive installed), and then insert the first Windows installation disk into the DVD drive. Skip to Step 6.
 - To install from remote ISO, log in to the server's Cisco IMC interface and continue with the next step.
- Step 4** Launch a Virtual KVM console window and click the **Virtual Media** tab.
 - a) Click **Add Image** and browse to select your remote Windows installation ISO file.
 - b) Check the check box in the **Mapped** column for the media that you just added, and then wait for mapping to complete.
- Step 5** Power cycle the server.
- Step 6** Press **F6** when you see the F6 prompt during bootup. The Boot Menu window opens.
- Step 7** On the Boot Manager window, choose the physical disk or virtual DVD and press **Enter**. The Windows installation begins when the image is booted.

- Step 8** Press **Enter** when you see the prompt, “Press any key to boot from CD.”
- Step 9** Observe the Windows installation process and respond to prompts in the wizard as required for your preferences and company standards.
- Step 10** When Windows prompts you with “Where do you want to install Windows,” install the drivers for embedded MegaRAID:
- Click **Load Driver**. You are prompted by a Load Driver dialog box to select the driver to be installed.
 - Connect the USB thumb drive that you prepared in Step 3 to the target server.
 - On the Windows Load Driver dialog, click **Browse**.
 - Use the dialog box to browse to the location of the drivers folder on the USB thumb drive, and then click **OK**.
Windows loads the drivers from the folder and when finished, the driver is listed under the prompt, “Select the driver to be installed.”
 - Click **Next** to install the drivers.

Updating Microsoft Windows Server Drivers

Procedure

- Click **Start**, point to **Settings**, and then click **Control Panel**.
- Double-click **System**, click the **Hardware** tab, and then click **Device Manager**. Device Manager starts.
- In **Device Manager**, double-click **SCSI and RAID Controllers**, right-click the device for which you are installing the driver, and then click **Properties**.
- On the **Driver** tab, click **Update Driver** to open the **Update Device Driver** wizard, and then follow the wizard instructions to update the driver.

Linux Drivers

Downloading the Driver Image File

See [Downloading the MegaSR Drivers, on page 60](#) for instructions on downloading the drivers. The Linux driver is included in the form of `dud-[driver version].img`, which is the boot image for the embedded MegaRAID stack.



Note The LSI MegaSR drivers that Cisco provides for Red Hat Linux and SUSE Linux are for the original GA versions of those distributions. The drivers do not support updates to those OS kernels.

Preparing Physical Thumb Drive for Linux

This topic describes how to prepare a physical Linux thumb drive from the driver image files.

This procedure requires a CD or DVD drive that you can use to burn the ISO image to disk; and a USB thumb drive.

Alternatively, you can mount the `dud.img` file as a virtual floppy disk, as described in the installation procedures. For RHEL and SLES, you can use a driver disk utility to create disk images from image files.

Procedure

Step 1 Download the Cisco UCS B-Series drivers ISO, as described in [Downloading the MegaSR Drivers, on page 60](#) and save it to your Linux system.

Step 2 Extract the `dud.img` or `dd.iso` driver file:

Note For RHEL 7.1 and later, there is no `dud.img` file--the driver is contained in a `dd.iso` file.

- a) Burn the Cisco UCS C-Series Drivers ISO image to a disc.
- b) Browse the contents of the drivers folders to the location of the embedded MegaRAID drivers:
/`<OS>`/Storage/Intel/C600-M5/
- c) Expand the Zip file, which contains the folder with the driver files.

Step 3 Copy the driver update disk image `dud-[driver version].img` (or `dd.iso`) to your Linux system.

Step 4 Insert a blank USB thumb drive into a port on your Linux system.

Step 5 Create a directory and mount the `dud.img` or `dd.iso` image to that directory:

Example:

```
mkdir <destination_folder>
mount -o loop <driver_image> <destination_folder>
```

Step 6 Copy the contents in the directory to your USB thumb drive.

Installing the Red Hat Enterprise Linux Driver

For the specific supported OS versions, see the [Hardware and Software Compatibility Matrix](#) for your server release.

This topic describes the fresh installation of the RHEL device driver on systems that have the embedded MegaRAID stack.



Note If you use an embedded RAID controller with Linux, both the pSATA and the sSATA controller must be set to `LSI SW RAID` mode.

Before you begin

Before you install this driver on the sSATA embedded controller, you must configure a RAID drive group.

To access the configuration utility, open the BIOS Setup Utility, go to the **Advanced** tab, and then choose the utility instance for the sSATA embedded controller: **LSI Software RAID Configuration Utility (sSATA)**.

Procedure

Step 1

Prepare the `dud.img` file using one of the following methods:

Note For RHEL 7.1 and later, there is no `dud.img` file--the driver is contained in a `dd.iso` file.

- To install from a physical disk, use the procedure in [Preparing Physical Thumb Drive for Linux, on page 62](#), then continue with Step 4.
- To install from a *virtual* disk, download the Cisco UCS B-Series drivers' ISO, as described in [Downloading the MegaSR Drivers, on page 60](#), then continue with the next step.

Step 2

Extract the `dud.img` (or `dd.iso`) file:

- a) Burn the Cisco UCS C-Series Drivers ISO image to a disk.
- b) Browse the contents of the drivers folders to the location of the embedded MegaRAID drivers:
`/<OS>/Storage/Intel/C600-M5/`
- c) Copy the `dud-<driver version>.img` (or `dd.iso`) file to a temporary location on your workstation.
- d) If you are using RHEL 7.x, rename the saved `dd.iso` to `dd.img`.

Note If you are using RHEL 7.x, renaming the `dd.iso` file to `dd.img` simplifies this procedure and saves time. The Cisco UCS virtual drive mapper can map only one `.iso` at a time, and only as a virtual CD/DVD. Renaming the file to `dd.img` allows you to mount the RHEL installation ISO as a virtual CD/DVD and the renamed `dd.img` as a virtual floppy disk or removable disk at the same time. This avoids the steps of unmounting and remounting the RHEL ISO when the `dd.iso` driver file is prompted for.

Step 3

Start the Linux driver installation using one of the following methods:

- To install from local media, connect an external USB CD/DVD drive to the server and then insert the first RHEL installation disk into the drive. Then continue with Step 5.
- To install from virtual disk, log in to the server's Cisco IMC interface. Then continue with the next step.

Step 4

Launch a Virtual KVM console window and click the **Virtual Media** tab.

- a) Click **Add Image** and browse to select your remote RHEL installation ISO image.

Note An `.iso` file can be mapped only as a virtual CD/DVD.

- b) Click **Add Image** again and browse to select your RHEL 6.x `dud.img` or the RHEL 7.x `dd.img` file that you renamed in step 2.

Note Map the `.img` file as a virtual floppy disk or virtual removable disk.

- c) Check the check boxes in the **Mapped** column for the media that you just added, then wait for mapping to complete.

Step 5

Power-cycle the target server.

Step 6

Press **F6** when you see the F6 prompt during bootup. The Boot Menu window opens.

Note Do not press Enter in the next step to start the installation. Instead, press **e** to edit installation parameters.

Step 7 On the Boot Menu window, use the arrow keys to select **Install Red Hat Enterprise Linux** and then press **e** to edit installation parameters.

Step 8 Append one of the following blacklist commands to the end of the line that begins with **linuxefi**:

- For RHEL 6.x (32- and 64-bit), type:

```
linux dd blacklist=iscsi blacklist=ahci nodmraid noprobe=<atadrive number>
```

Note The noprobe values depend on the number of drives. For example, to install RHEL 6.x on a RAID 5 configuration with three drives, type:

```
Linux dd blacklist=iscsi blacklist=ahci nodmraid noprobe=ata1 noprobe=ata2
```

- For RHEL 7.x (32- and 64-bit), type:

```
linux dd modprobe.blacklist=ahci nodmraid
```

Step 9 **Optional:** To see full, verbose installation status steps during installation, delete the **Quiet** parameter from the line.

Step 10 On the Boot Menu window, press **Ctrl+x** to start the interactive installation.

Step 11 Below **Driver disk device selection**, select the option to install your driver .img file. (Type **r** to refresh the list if it is not populated.)

Note The installer recognizes the driver file as an .iso file, even though you renamed it to dd.img for mapping.

Type the number of the driver device ISO in the list. Do *not* select the RHEL ISO image. In the following example, type **6** to select device sdb:

```
5) sr0 iso9660 RHEL-7.6\x20Server.x
6) sdb iso9660 CDROM
# to select, 'r' - refresh, or 'c' -continue: 6
```

The installer reads the driver file and lists the drivers.

Step 12 Under **Select drivers to install**, type the number of the line that lists the megasr driver. In the following example, type **1**:

```
1) [ ] /media/DD-1/rpms/x86_61/kmod-megasr-18.01.2010.1107_e17.6-1.x86_61.rpm
# to toggle selection, or 'c' -continue: 1
```

Your selection is displayed with an X in brackets.

```
1) [X] /media/DD-1/rpms/x86_61/kmod-megasr-18.01.2010.1107_e17.6-1.x86_61.rpm
```

Step 13 Type **c** to continue.

Step 14 Follow the RHEL installation wizard to complete the installation.

Step 15 When the wizard's Installation Destination screen is displayed, ensure that **LSI MegaSR** is listed as the selection. If it is not listed, the driver did not load successfully. In that case, select **Rescan Disc**.

Step 16 After the installation completes, reboot the target server.

Installing the SUSE Linux Enterprise Server Driver

For the specific supported OS versions, see the [Hardware and Software Compatibility Matrix](#) for your server release.

This topic describes the fresh installation of the SLES driver on systems that have the embedded MegaRAID stack.



Note If you use an embedded RAID controller with Linux, both the pSATA and the sSATA controller must be set to `LSI SW RAID` mode.

Before you begin

Before you install this driver on the sSATA embedded controller, you must configure a RAID drive group.

To access the configuration utility, open the BIOS Setup Utility, go to the **Advanced** tab, and then choose the utility instance for the sSATA embedded controller: **LSI Software RAID Configuration Utility (sSATA)**.

Procedure

- Step 1** Prepare the `dud.img` file using one of the following methods:
- To install from a physical disk, use the procedure in [Preparing Physical Thumb Drive for Linux, on page 62](#), then continue with Step 4.
 - To install from a *virtual* disk, download the Cisco UCS B-Series drivers' ISO, as described in [Downloading the MegaSR Drivers, on page 60](#), then continue with the next step.
- Step 2** Extract the `dud.img` file that contains the driver:
- a) Burn the ISO image to a disk.
 - b) Browse the contents of the drivers folders to the location of the embedded MegaRAID drivers:
`/<OS>/Storage/Intel/C600-M5/...`
 - c) Within the SLES folder for your version, the `dud-<driver version>.img` file is packaged in a compressed `.gz` file. Extract the `.img` file from the `.gz` file.
 - d) Copy the `dud-<driver version>.img` file to a temporary location on your workstation.
- Step 3** Start the Linux driver installation using one of the following methods:
- To install from local media, connect an external USB DVD drive to the server and then insert the first SLES installation disk into the drive. Then continue with Step 5.
 - To install from remote ISO, log in to the server's Cisco IMC interface. Then continue with the next step.
- Step 4** Launch a Virtual KVM console window and click the **Virtual Media** tab.
- a) Click **Add Image** and browse to select your remote SLES installation ISO file.
 - b) Click **Add Image** again and browse to select your `dud-<driver version>.img` file.
 - c) Check the check boxes in the **Mapped** column for the media that you just added, then wait for mapping to complete.
- Step 5** Power-cycle the target server.

- Step 6** Press **F6** when you see the F6 prompt during bootup. The Boot Menu window opens.
- Step 7** On the Boot Manager window, select the physical or virtual SLES installation ISO and press **Enter**.
The SLES installation begins when the image is booted.
- Step 8** When the first SLES screen appears, select **Installation**.
- Step 9** Press **e** to edit installation parameters.
- Step 10** Append the following parameter to the end of the line that begins with **linuxefi**:
`brokenmodules=ahci`
- Step 11** **Optional:** To see detailed status information during the installation, add the following parameter to the line that begins with **linuxefi**:
`splash=verbose`
- Step 12** Press **Ctrl+x** to start the installation.
The installation proceeds. The installer finds the LSI driver automatically in the `dud-<driver version>.img` file that you provided. With verbose status messages, you see the driver being installed when `LSI MegaRAID SW RAID Module` is listed.
- Step 13** Follow the SLES installation wizard to complete the installation. Verify installation of the driver when you reach the **Suggested Partitioning** screen:
a) On the **Suggested Partitioning** screen, select **Expert Partitioner**.
b) Navigate to **Linux > Hard disks** and verify that there is a device listed for the `LSI - LSI MegaSR` driver. The device might be listed as a type other than `sda`. For example:
`dev/sdd: LSI - LSI MegaSR`
If no device is listed, the driver did not install properly. In that case, repeat the steps above.
- Step 14** When installation is complete, reboot the target server.

For More RAID Utility Information

The Broadcom utilities have help documentation for more information about using the utilities.

- For embedded software MegaRAID and the utility that is accessed via the server BIOS (refer to Chapter 4)—[Broadcom Embedded MegaRAID Software User Guide, March 2018](#).

Replacing a Boot-Optimized M.2 RAID Controller Module

The Cisco Boot-Optimized M.2 RAID Controller module connects to the mini-storage module socket on the motherboard. It includes slots for two SATA M.2 drives, plus an integrated 6-Gbps SATA RAID controller that can control the SATA M.2 drives in a RAID 1 array.

Cisco Boot-Optimized M.2 RAID Controller Considerations

Review the following considerations:



Note The Cisco Boot-Optimized M.2 RAID Controller is not supported when the server is used as a compute-only node in Cisco HyperFlex configurations.

- The minimum version of Cisco IMC and Cisco UCS Manager that support this controller is 4.0(4) and later.
- This controller supports RAID 1 (single volume) and JBOD mode.



Note Do not use the server's embedded SW MegaRAID controller to configure RAID settings when using this controller module. Instead, you can use the following interfaces:

- Cisco IMC 4.0(4a) and later
 - BIOS HII utility, BIOS 4.0(4a) and later
 - Cisco UCS Manager 4.0(4a) and later (UCS Manager-integrated servers)
-
- A SATA M.2 drive in slot 1 (the top) is the first SATA device; a SATA M.2 drive in slot 2 (the underside) is the second SATA device.
 - The name of the controller in the software is MStOR-RAID.
 - A drive in Slot 1 is mapped as drive 253; a drive in slot 2 is mapped as drive 254.
 - When using RAID, we recommend that both SATA M.2 drives are the same capacity. If different capacities are used, the smaller capacity of the two drives is used to create a volume and the rest of the drive space is unusable.

JBOD mode supports mixed capacity SATA M.2 drives.
 - Hot-plug replacement is *not* supported. The server must be powered off.
 - Monitoring of the controller and installed SATA M.2 drives can be done using Cisco IMC and Cisco UCS Manager. They can also be monitored using other utilities such as UEFI HII, PMCLI, XMLAPI, and Redfish.
 - To upgrade and manage firmware for the controller and M.2 drives by using Cisco UCS Manager, refer to the [Cisco UCS Manager Firmware Management Guide](#).
 - The SATA M.2 drives can boot in UEFI mode only. Legacy boot mode is not supported.
 - If you replace a single SATA M.2 drive that was part of a RAID volume, rebuild of the volume is auto-initiated after the user accepts the prompt to import the configuration. If you replace both drives of a volume, you must create a RAID volume and manually reinstall any OS.
 - We recommend that you erase drive contents before creating volumes on used drives from another server. The configuration utility in the server BIOS includes a SATA secure-erase function.
 - The server BIOS includes a configuration utility specific to this controller that you can use to create and delete RAID volumes, view controller properties, and erase the physical drive contents. Access the utility

by pressing **F2** when prompted during server boot. Then navigate to **Advanced > Cisco Boot Optimized M.2 RAID Controller**.

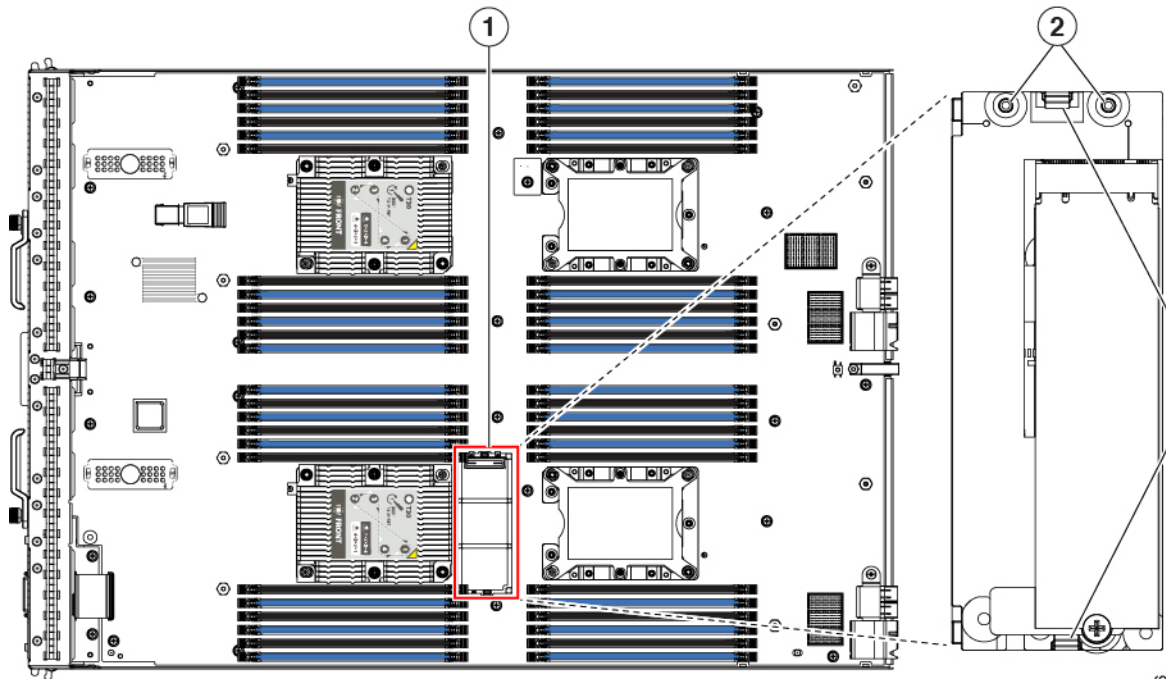
Replacing a Cisco Boot-Optimized M.2 RAID Controller

This topic describes how to remove and replace a Cisco Boot-Optimized M.2 RAID Controller. The controller board has one M.2 socket on its top (Slot 1) and one M.2 socket on its underside (Slot 2).

Procedure

- Step 1** Remove the controller from the server:
- Decommission, power off, and remove the blade server from the chassis.
 - Remove the top cover from the server as described in [Removing a Blade Server Cover, on page 14](#).
 - Press out on the securing clips to disengage the controller from the socket.
 - Pull straight up on the controller to remove it.

Figure 38: Cisco Boot-Optimized M.2 RAID Controller on Motherboard



307206

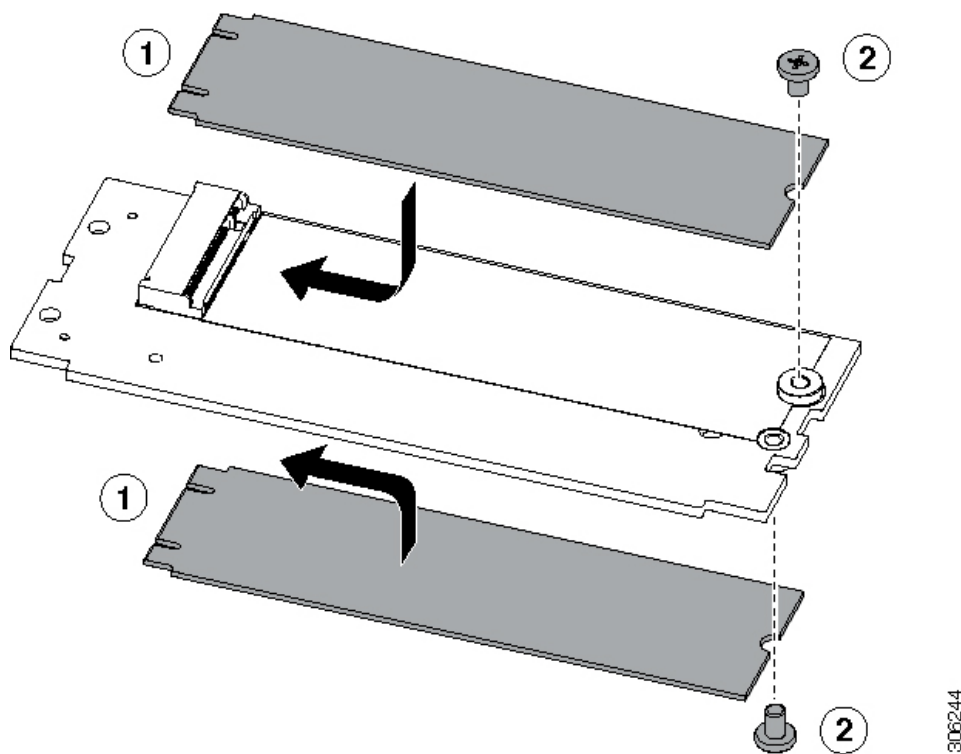
1	Location of socket on motherboard	3	Securing clips
2	Alignment pegs	-	

- Step 2** If you are transferring SATA M.2 drives from the old controller to the replacement controller, do that before installing the replacement controller:

Note Any previously configured volume and data on the drives are preserved when the M.2 drives are transferred to the new controller. The system will boot the existing OS that is installed on the drives.

- a) Use a #1 Phillips-head screwdriver to remove the single screw that secures the M.2 drive to the carrier.
- b) Lift the M.2 drive from its socket on the carrier.
- c) Position the replacement M.2 drive over the socket on the controller board.
- d) Angle the M.2 drive downward and insert the connector-end into the socket on the carrier. The M.2 drive's label must face up.
- e) Press the M.2 drive flat against the carrier.
- f) Install the single screw that secures the end of the M.2 SSD to the carrier.
- g) Turn the controller over and install the second M.2 drive.

Figure 39: Cisco Boot-Optimized M.2 RAID Controller, Showing M.2 Drive Installation



Step 3 Install the controller to its socket on the motherboard:

- a) Position the controller over the socket, with the controller's connector facing down and at the same end as the motherboard socket. Two alignment pegs must match with two holes on the controller.
- b) Gently push down the socket end of the controller so that the two pegs go through the two holes on the carrier.
- c) Push down on the controller so that the securing clips click over it at both ends.

Step 4 Replace the top cover to the server.

Step 5 Return the blade server to the chassis and allow it to be automatically reacknowledged, reassociated, and recommissioned.

Recycling the PCB Assembly (PCBA)

Each blade server has a PCBA that is connected to the blade server's faceplate and sheet metal tray. You must disconnect the PCBA from the blade server's faceplate and tray to recycle the PCBA. Each blade server is attached to the faceplate and tray by the following screws:

- Faceplate: Two M2x0.4mm screws.
- Tray: 33 M3x0.5mm screws.

You will need to recycle the PCBA for each blade server.

Before you begin



Note **For Recyclers Only!** This procedure is not a standard field-service option. This procedure is for recyclers who will be reclaiming the electronics for proper disposal to comply with local eco design and e-waste regulations.

To remove the printed circuit board assembly (PCBA), the following requirements must be met:

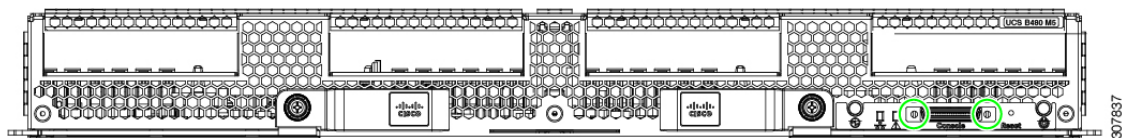
- The server must be disconnected from facility power.
- The server must be removed from the equipment rack.
- The server's top cover must be removed. See [Removing a Blade Server Cover, on page 14](#).
- If you have not already removed the TPM module, do so now. See [Removing the Trusted Platform Module \(TPM\), on page 54](#).

Procedure

Step 1 Using a screwdriver, rotate each of the faceplate screws counter clockwise until it disengages.

The following figure shows the location of these screws.

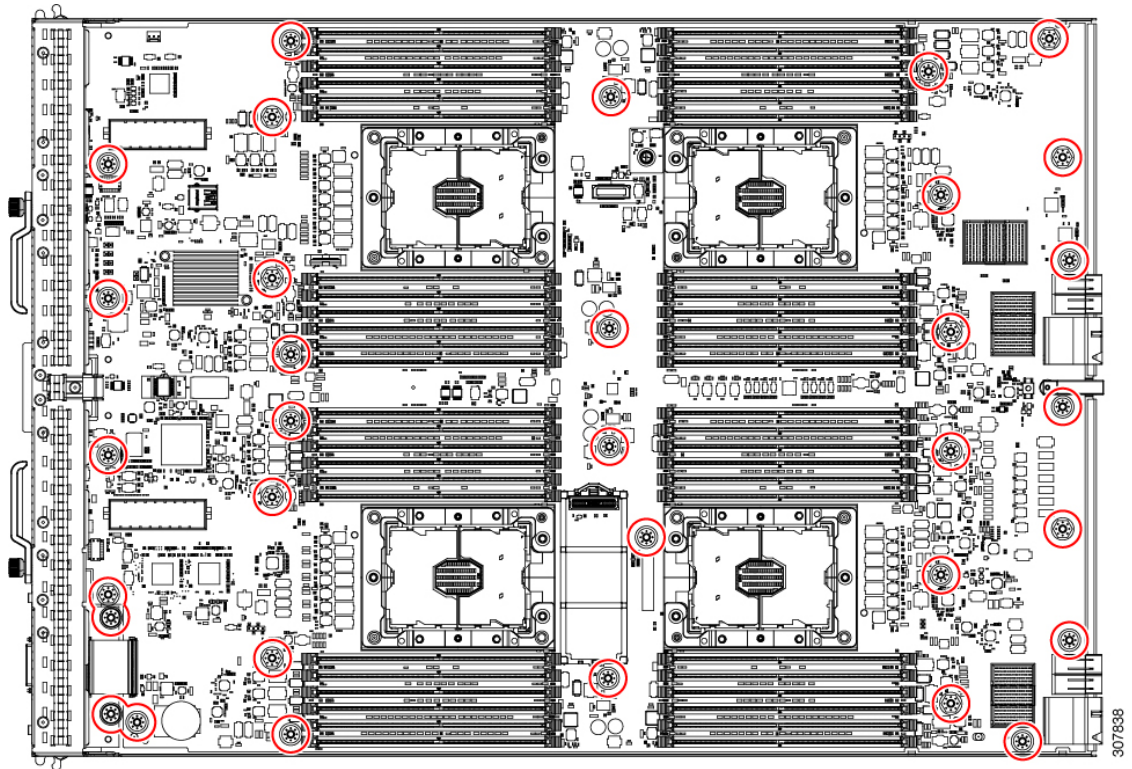
Figure 40: Location of PCBA Mounting Screws on Faceplate



Step 2 Using a screwdriver, rotate each of the mounting screws counter-clockwise until it disengages.

The following figure shows the location of these screws.

Figure 41: Location of PCBA Mounting Screws on Blade Server Tray



Step 3 Detach the PCBA from the faceplate and tray and dispose of all parts properly.



CHAPTER 4

Technical Specifications

This chapter contains the following sections:

- [Physical Specifications for the UCS B480 Blade Server, on page 73](#)

Physical Specifications for the UCS B480 Blade Server

Specification	Value
Height	1.95 inches (50 mm)
Width	16.50 inches (419.1 mm)
Depth	24.4 inches (620 mm)
Weight	<p>The weight depends on the components installed.</p> <ul style="list-style-type: none">• Base server weight: 17.7 lbs (8.0 kg) (No CPUs, no memory, no adapter cards, no USB, no SD cards, no HDDs, 4 HDD fillers, no TPM)• Minimally configured server weight: 21.09 lbs (9.56 kg) (2 CPUs, 2 heat sinks, 2 DIMMs, mLOM, no adapter cards, 4 HDD fillers)• Fully configured server weight: 30.7 lbs (13.9 kg) (4 CPUs, 4 heat sinks, 48 DIMMs, mLOM, 2 adapter cards, 4 HDDs)



CHAPTER 5

NVIDIA Licensing Information

This chapter contains the following sections:

- [NVIDIA GRID License Server Overview, on page 75](#)
- [Registering Your Product Activation Keys with NVIDIA, on page 76](#)
- [Downloading the GRID Software Suite, on page 77](#)
- [Installing NVIDIA GRID License Server Software, on page 77](#)
- [Installing GRID Licenses From the NVIDIA Licensing Portal to the License Server, on page 79](#)
- [Managing GRID Licenses, on page 81](#)
- [Installing Drivers to Support the NVIDIA GPU Cards, on page 82](#)

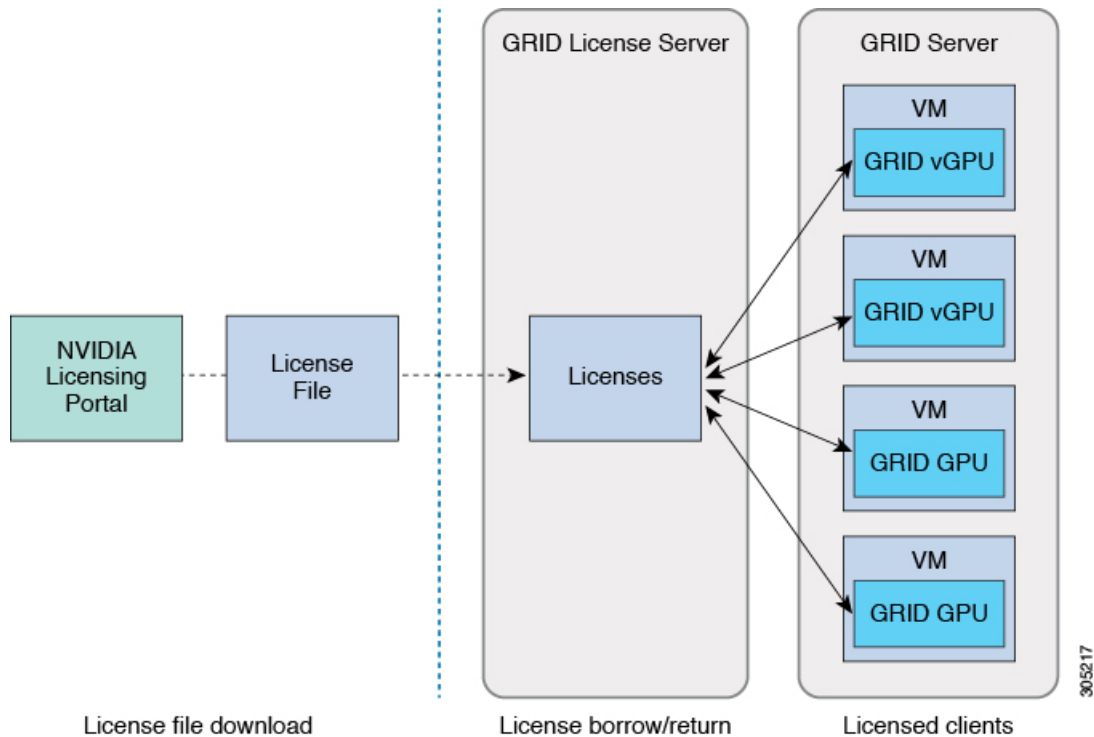
NVIDIA GRID License Server Overview

The NVIDIA Tesla P6 GPU combines Tesla and GRID functionality when you enable the licensed GRID features *GRID vGPU* and *GRID Virtual Workstation*. You enable these features during OS boot by borrowing a software license that is served over the network from the NVIDIA GRID License Server virtual appliance. The license is returned to the GRID License Server when the OS shuts down.

The NVIDIA Tesla P6 GPU has dual personality. It can work in Compute (Tesla) and GRID mode. Only GRID mode needs a license.

You obtain the licenses that are served by the GRID License Server from the NVIDIA Licensing Portal as downloadable license files, which you install into the GRID License Server via its management interface. See the following figure.

Figure 42: GRID Licensing Architecture



There are three editions of GRID licenses that enable three different classes of GRID features. The GRID software automatically selects the license edition based on the features that you are using. See the following table.

Table 5: GRID Licensing Editions

GRID License Edition	GRID Features
GRID Virtual GPU (vGPU)	Virtual GPUs for business desktop computing
GRID Virtual Workstation	Virtual GPUs for mid-range workstation computing
GRID Virtual Workstation - Extended	Virtual GPUs for high-end workstation computing Workstation graphics on GPU pass-through

Registering Your Product Activation Keys with NVIDIA

After your order is processed, NVIDIA sends you a Welcome email that contains your product activation keys (PAKs) and a list of the types and quantities of licenses that you purchased.

Procedure

-
- Step 1** Select the **Log In** link, or the **Register** link if you do not already have an account.

- The NVIDIA Software Licensing Center > License Key Registration dialog opens.
- Step 2** Complete the License Key Registration form and then click **Submit My Registration Information**. The NVIDIA Software Licensing Center > Product Information Software Dialog opens.
- Step 3** If you have additional PAKs, click **Register Additional Keys**. For each additional key, complete the form on the License Key Registration dialog, and then click **Submit My Registration Information**.
- Step 4** Agree to the terms and conditions and set a password when prompted.
-

Downloading the GRID Software Suite

Procedure

- Step 1** Return to the NVIDIA Software Licensing Center > Product Information Software dialog box.
- Step 2** Click **Current Releases**.
- Step 3** Click the **NVIDIA GRID** link to access the **Product Download** dialog. This dialog includes download links for:
- The NVIDIA License Manager software
 - The `gpumodeswitch` utility
 - The host driver software
- Step 4** Use the links to download the software.
-

Installing NVIDIA GRID License Server Software

For full instructions and troubleshooting information, see the *NVIDIA GRID License Server User Guide*. Also see the *NVIDIA GRID License Server Release Notes* for the latest information about your release. Both documents are available at the following URL:

<http://www.nvidia.com>

Platform Requirements for NVIDIA GRID License Server

- The hosting platform can be a physical or a virtual machine. NVIDIA recommends using a host that is dedicated to running only the License Server.
- The hosting platform must run a supported Windows OS.
- The hosting platform must have a constant IP address.
- The hosting platform must have at least one constant Ethernet MAC address.
- The hosting platform's date and time must be set accurately.

Installing on Windows

Before you begin

The License Server requires a Java Runtime Environment and an Apache Tomcat installation. Apache Tomcat is installed when you use the NVIDIA installation wizard for Windows.

Procedure

- Step 1** Download and install the latest Java 32-bit Runtime Environment from <https://www.oracle.com/downloads/index.html>.
- Note** Install the 32-bit Java Runtime Environment, regardless of whether your platform is Windows 32-bit or 64-bit.
- Step 2** Create a server interface.
- In the **NVIDIA Software Licensing Center** dialog box, click **Grid Licensing > Create License Server**.
 - In the **Create Server** dialog box, fill in your desired server details.
 - Save the .bin file that is generated to your license server for installation.
- Step 3** Unzip the NVIDIA License Server installer Zip file that you downloaded previously and run `setup.exe`.
- Step 4** Accept the EULA for the NVIDIA License Server software and the Apache Tomcat software. Tomcat is installed automatically during the License Server installation.
- Step 5** Use the installer wizard to step through the installation.
- Note** In the **Choose Firewall Options** dialog box, select the ports to be opened in the firewall. NVIDIA recommends that you use the default setting, which opens port 7070 but leaves port 8080 closed.
- Step 6** To verify the installation, open a web browser on the License Server host and connect to the URL `http://localhost:8080/licserver`. If the installation was successful, you see the NVIDIA Licenses Client Manager interface.
-

Installing on Linux

Before you begin

The License Server requires a Java Runtime Environment and an Apache Tomcat installation. Use the following commands to install both separately before installing the License Server on Linux.

Procedure

- Step 1** Verify that Java was installed with your Linux installation:
- ```
java -version
```
- If a Java version does not display, use your Linux package manager to install Java:

```
sudo yum install java
```

**Step 2** Use your Linux package manager to install the Tomcat and Tomcat webapp packages.

a) Install Tomcat:

```
sudo yum install tomcat
```

b) Enable the Tomcat service for automatic startup on boot:

```
sudo systemctl enable tomcat.service
```

c) Start the Tomcat service:

```
sudo systemctl start tomcat.service
```

d) To verify that the Tomcat service is operational, open a web browser on the License Server host and connect to the URL <http://localhost:8080>. If the installation was successful, you see the Tomcat webapp.

**Step 3** Install the License Server.

a) Unpack the License Server tar file:

```
tar xzf NVIDIA-linux-2015.09-0001.tgz
```

b) Run the unpacked setup binary as root:

```
sudo ./setup.bin
```

c) Accept the EULA and then continue with the installation wizard to finish the installation.

**Note** In the **Choose Firewall options** dialog, select the ports to be opened in the firewall. NVIDIA recommends that you use the default setting, which opens port 7070 but leaves port 8080 closed.

**Step 4** To verify the installation, open a web browser on the License Server host and connect to the URL <http://localhost:8080/licserver>. If the installation was successful, you see the NVIDIA License Client Manager interface.

---

# Installing GRID Licenses From the NVIDIA Licensing Portal to the License Server

## Accessing the GRID License Server Management Interface

Open a web browser on the License Server host and access the URL <http://localhost:8080/licserver>.

If you configure the License Server host's firewall to permit remote access to the License Server, the management interface is accessible from remote machines at the URL <http://localhost:8080/licserver>.

## Reading Your License Server's MAC Address

Your License Server's Ethernet MAC address is used as an identifier when registering the License Server with NVIDIA's Licensing Portal.

### Procedure

---

- Step 1** Access the GRID License Server Management Interface in a browser.
- Step 2** In the left-side **License Server** panel, select **Configuration**.
- Step 3** In the **License Server Configuration** panel, from the **Server host ID** pull-down menu, select your License Server's Ethernet MAC address.
- Note** It is important to use the same Ethernet ID consistently to identify the server when generating licenses on NVIDIA's Licensing Portal. NVIDIA recommends that you select one entry for a primary, non-removable Ethernet interface on the platform.
- 

## Installing Licenses From the Licensing Portal

### Procedure

---

- Step 1** Access the GRID License Server Management Interface in a browser.
- Step 2** In the left-side **License Server** panel, select **Configuration**.
- Step 3** From the **License Server Configuration** menu, click **Choose File**.
- Step 4** Browse to the license .bin file that you generated earlier and want to install, and click **Open**.
- Step 5** Click **Upload**.  
The license file installs on your License Server. When the installation is complete, you see the confirmation message, "Successfully applied license file to license server."
- 

## Viewing Available Licenses

Use the following procedure to view the licenses that are installed and available and their properties.

### Procedure

---

- Step 1** Access the GRID License Server Management Interface in a browser.
- Step 2** In the left-side **License Server** panel, select **Licensed Feature Usage**.
- Step 3** Click a feature in the **Feature** column to see detailed information about the current usage of that feature.
- 

## Viewing Current License Usage

Use the following procedure to view information about that licenses that are currently in-use and borrowed from the server.

### Procedure

---

- Step 1** Access the GRID License Server Management Interface in a browser.
- Step 2** In the left-side **License Server** panel, select **Licensed Clients**.
- Step 3** To view detailed information about a single licensed client, click its **Client ID** in the list.
- 

## Managing GRID Licenses

Features that require GRID licensing run at reduced capability until a GRID license is acquired.

### Acquiring a GRID License on Windows

To acquire a GRID license on Windows, use the following procedure.

#### Procedure

---

- Step 1** Open the NVIDIA Control Panel using one of the following methods:
- Right-click the Windows desktop and select **NVIDIA Control Panel** from the menu.
  - Open the Windows Control Panel and double-click the **NVIDIA Control Panel** icon.
- Step 2** In the **NVIDIA Control Panel** left pane under **Licensing**, select **Manage License**. The **Manage License** task pane opens and shows the current license edition being used. The GRID software automatically selects the license edition based on the feature that you are using. The default is **Tesla (unlicensed)**.
- Step 3** If you want to acquire a license for GRID Virtual Workstation, under **License Edition**, select **GRID Virtual Workstation**.
- Step 4** In the **License Server** field, enter the address of your local GRID License Server. The address can be a domain name or an IP address.
- Step 5** In the **Port Number** field, enter your port number or leave it set to the default used by the server, which is 7070.
- Step 6** Select **Apply**. The system requests the appropriate license edition from your configured License Server. After a license is successfully acquired, the features of that license edition are enabled.

**Note** After you configure licensing settings in the NVIDIA Control Panel, the settings persist across reboots.

---

## Acquiring a GRID License on Linux

To acquire a GRID license on Linux, use the following procedure.

### Procedure

---

**Step 1** Edit the configuration file:

```
sudo vi /etc/nvidia/gridd.conf
```

**Step 2** Edit the `ServerUrl` line with the address of your local GRID License Server.

The address can be a domain name or an IP address. See the Sample Configuration File.

**Step 3** Append the port number (default **7070**) to the end of the address with a colon. See the Sample Configuration File.

**Step 4** Edit the `FeatureType` line with an integer for the license type. See the Sample Configuration File.

- GRID vGPU = 1
- GRID Virtual Workstation = 2

**Step 5** Restart the `nvidia-gridd` service:

```
sudo service nvidia-gridd restart
```

The service automatically acquires the license edition that you specified in the `FeatureType` line. You can confirm this in `/var/log/messages`.

**Note** After you configure licensing settings in `gridd.conf`, the settings persist across reboots.

### Sample Configuration File

```
/etc/nvidia/gridd.conf - Configuration file for NVIDIA Grid Daemon
Description: Set License Server URL
Data type: string
Format: "<address>:<port>"
Server URL=10.31.20.45:7070

Description: set Feature to be enabled
Data type: integer
Possible values:
1 => for GRID vGPU
2 => for GRID Virtual Workstation
FeatureType=1
```

---

## Installing Drivers to Support the NVIDIA GPU Cards

After you install the hardware, you must update to the correct level of server BIOS, activate the BIOS firmware, and then install NVIDIA drivers and other software in this order:

# 1. Updating the Server BIOS Firmware

Install the latest Cisco server BIOS for your blade server by using Cisco UCS Manager.



**Note** You must do this procedure before you update the NVIDIA drivers.



**Caution** Do not remove the hardware that contains the endpoint or perform any maintenance on it until the update process completes. If the hardware is removed or otherwise unavailable due to maintenance, the firmware update fails. This failure might corrupt the backup partition. You cannot update the firmware on an endpoint with a corrupted backup partition.

## Procedure

- Step 1** In the **Navigation** pane, click **Equipment**.
- Step 2** On the **Equipment** tab, expand **Equipment > Chassis > Chassis Number > Servers**.
- Step 3** Click the **Name** of the server for which you want to update the BIOS firmware.
- Step 4** On the **Properties** page in the **Inventory** tab, click **Motherboard**.
- Step 5** In the **Actions** area, click **Update BIOS Firmware**.
- Step 6** In the **Update Firmware** dialog box, do the following:
  - a) From the **Firmware Version** drop-down list, select the firmware version to which you want to update the endpoint.
  - b) Click **OK**.

Cisco UCS Manager copies the selected firmware package to the backup memory slot, where it remains until you activate it.
- Step 7** (Optional) Monitor the status of the update in the **Update Status** field.

The update process can take several minutes. Do not activate the firmware until the firmware package you selected displays in the **Backup Version** field in the **BIOS** area of the **Inventory** tab.

## What to do next

Activate the server BIOS firmware.

# 2. Activating the Server BIOS Firmware

## Procedure

- Step 1** In the **Navigation** pane, click **Equipment**.
- Step 2** On the **Equipment** tab, expand **Equipment > Chassis > Chassis Number > Servers**.

**Step 3** Click the **Name** of the server for which you want to activate the BIOS firmware.

**Step 4** On the **Properties** page in the **Inventory** tab, click **Motherboard**.

**Step 5** In the **Actions** area, click **Activate BIOS Firmware**.

**Step 6** In the **Activate Firmware** dialog box, do the following:

- a) Select the appropriate server BIOS version from the **Version To Be Activated** drop-down list.
- b) If you want to set only the start-up version and not change the version running on the server, check **Set Startup Version Only**.

If you configure **Set Startup Version Only**, the activated firmware moves into the pending-next-reboot state and the server is not immediately rebooted. The activated firmware does not become the running version of firmware until the server is rebooted.

- c) Click **OK**.

---

#### What to do next

Update the NVIDIA drivers.

## 3. Updating the NVIDIA Drivers

After you update the server BIOS, you can install NVIDIA drivers to your hypervisor virtual machine.

#### Procedure

---

**Step 1** Install your hypervisor software on a computer. Refer to your hypervisor documentation for the installation instructions.

**Step 2** Create a virtual machine in your hypervisor. Refer to your hypervisor documentation for instructions.

**Step 3** Install the NVIDIA drivers to the virtual machine. Download the drivers:

- NVIDIA Enterprise Portal for GRID hypervisor downloads (requires NVIDIA login):  
<https://nvidia.flexnetoperations.com/>
- NVIDIA public driver area: <http://www.nvidia.com/Download/index.aspx>

**Step 4** Restart the server.

**Step 5** Check that the virtual machine is able to recognize the NVIDIA card. In Windows, use the **Device Manager** and look under **Display Adapters**.

---