



NVIDIA DGX A100 Service Manual

NVIDIA Corporation

Apr 27, 2024

Contents

1	Introduction	1
1.1	Customer-replaceable Components	1
1.2	Recommended Tools	2
1.3	Customer Support	2
2	Front Fan Module Replacement	3
2.1	Front Fan Module Replacement Overview	3
2.2	Identifying the Failed Fan Module	3
2.2.1	Viewing the Fan Module LED	4
2.2.2	Using the BMC Dashboard and NVSM	4
2.3	Replacing and Returning the Front Fan Module	5
3	Power Supply Replacement	7
3.1	Power Supply Replacement Overview	7
3.2	Identifying the Failed Power Supply	7
3.2.1	Identifying the Failed Power Supply from the Back	7
3.2.2	Identifying the Failed Power Supply from the Console	8
3.2.3	Determining the Manufacturer	9
3.3	Replacing the Power Supply	10
4	Motherboard Tray - Accessing in Place	11
4.1	Accessing the Motherboard Tray	11
4.2	Replacing the Motherboard Tray	13
5	Motherboard Tray - Removal and Installation	17
5.1	Removing the Motherboard Tray	17
5.2	Reinstalling the Motherboard Tray	21
6	U.2 NVMe Cache Drive Upgrade from 4 to 8	25
6.1	U.2 NVMe Cache Drive Upgrade Overview	25
6.2	Determining the NVMe Drive to Order	25
6.3	Installing the Optional NVMe Drives	26
7	U.2 NVMe Cache Drive Replacement	29
7.1	U.2 NVMe Cache Drive Replacement Overview	29
7.2	Identifying the Failed U.2 NVMe	29
7.2.1	Identifying the Failed NVMe from the Front	29
7.2.2	Identifying the Failed NVMe from the Console	30
7.2.3	Identifying the NVMe Manufacturer and Model	31
7.3	Replacing the U.2 NVMe Drive	31
8	U.2 NVMe Cache Drive Upgrade to 7.68 TB Drives	33
8.1	U.2 NVMe Cache Drive 7.68 TB Upgrade Overview	33
8.2	Installing the 7.68 TB NVMe Drives	33

9 U.2 NVMe Cache Drive Post-Installation Tasks	37
9.1 Recreating the Cache RAID 0 Volume	37
9.2 Returning the NVMe Drive	38
10 M.2 NVMe Boot Drive Replacement	39
10.1 M.2 NVMe Boot Drive Replacement Overview	39
10.2 Identifying the Failed M.2 NVMe	39
10.3 Replacing the M.2 NVMe Drive	40
10.4 Rebuilding the Boot Drive RAID 1 Volume	44
10.5 Returning the NVMe Drive	45
11 M.2 Boot Drive Riser Assembly Replacement	47
11.1 M.2 Boot Drive Riser Assembly Replacement Overview	47
11.2 Determining a Failed M.2 NVMe Riser Assembly	48
11.3 Replacing the M.2 NVMe Riser Assembly	48
11.4 Returning the Riser Assembly	50
12 DIMM Replacement	51
12.1 DIMM Replacement Overview	51
12.2 Identifying the Failed DIMM	51
12.3 Replacing the DIMM	52
13 DIMM Upgrade	57
13.1 DIMM Upgrade Overview	57
13.2 Upgrading the DIMM	57
14 Network Card Replacement	63
14.1 Network Card Replacement Overview	63
14.2 Identifying the Failed Network Card	63
14.3 Replacing the Vertical Network Card	64
14.4 Replacing the Horizontal Network Card	65
15 Adding the Optional Dual-port Horizontal Network Card	71
15.1 Dual-port Network Card Upgrade Overview	71
15.2 Adding the Horizontal Network Card	72
16 Updating the Mellanox Network Card Firmware	77
17 Front Console Board Replacement	79
17.1 Front Console Board Replacement Overview	79
17.2 Replacing the Front Console Board	79
18 Motherboard Tray Battery Replacement	83
18.1 Motherboard Tray Battery Replacement Overview	83
18.2 Replacing the Motherboard Tray Battery	84
19 Trusted Platform Module Replacement	89
19.1 Trusted Platform Module Replacement Overview	89
19.2 Replacing the Trusted Platform Module	89
20 Removing and Attaching the Bezel	93
21 Installing the Rack Mount Kit	95
21.1 Installing the Rails	95
21.2 Installing the Cage Nuts	97

22 Third-Party License Notices	99
22.1 Micron msecli	99
22.2 Mellanox (OFED)	100
23 Notices	101
23.1 Notice	101
23.2 Trademarks	102

Chapter 1. Introduction

This document contains instructions for replacing NVIDIA DGX™ A100 system components. Be sure to familiarize yourself with the NVIDIA Terms & Conditions documents before attempting to perform any modification or repair to the DGX A100 system. These Terms & Conditions for the DGX A100 system can be found through the [NVIDIA DGX Systems Support](#) page.

Contact NVIDIA Enterprise Support to obtain an RMA number for any system or component that needs to be returned for repair or replacement. When replacing a component, use **only** the replacement supplied to you by NVIDIA.

1.1. Customer-replaceable Components

List of customer-replaceable components in the NVIDIA DGX A100.

You can obtain the following components for replacement in your data center.

- ▶ Fan Modules
- ▶ Power Supplies
- ▶ Cache (U.2) NVMe Drives
- ▶ Boot (M.2) NVMe Drive
- ▶ Boot (M.2) Riser Assembly
- ▶ DIMMs
- ▶ Network Card (vertical, single or dual port)
- ▶ Network Card (horizontal, dual port)
- ▶ Front Console Board
- ▶ Battery
- ▶ Trusted Platform Module (TPM)
- ▶ Bezel
- ▶ Rack Mount Kit

Contact NVIDIA Enterprise Support for replacement instructions and guidance for specific components if those instructions are not included in this document.

1.2. Recommended Tools

List of recommended tools needed to service the NVIDIA DGX A100.

- ▶ Laptop
- ▶ USB key with tools and drivers
- ▶ USB key imaged with the DGX Server OS ISO
- ▶ Screwdrivers (Phillips #1 and #2, small flat head)
- ▶ KVM Crash Cart
- ▶ Anti-static wrist strap
- ▶ Masking tape or label maker
- ▶ Tie wraps or VELCRO® Brand for cable management
- ▶ Box cutter
- ▶ Black Permanent Marker
- ▶ Packing materials

1.3. Customer Support

Contact NVIDIA Enterprise Support for assistance in reporting, troubleshooting, or diagnosing problems with your DGX A100 system. Also contact NVIDIA Enterprise Support for assistance in installing or moving the DGX A100 system.

For details on how to obtain support, visit the NVIDIA Enterprise Support web site (<https://www.nvidia.com/en-us/support/enterprise/>).

Chapter 2. Front Fan Module Replacement

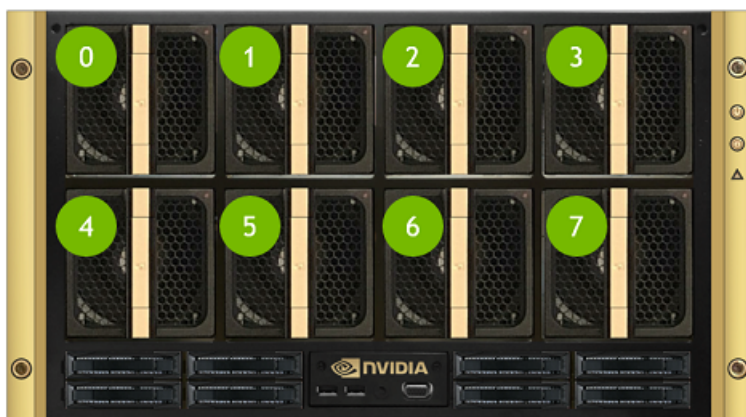
2.1. Front Fan Module Replacement Overview

This is a high-level overview of the steps needed to replace the front fan modules.

1. Identify the failed front fan module through the BMC or the fan module LED and submit a service ticket to NVIDIA Enterprise Support.
2. Get a replacement from NVIDIA Enterprise Support.
3. Remove the failed fan module using the fan numbering diagram as a reference.
4. Insert the new fan module.
5. Confirm the new fan module is working correctly through the BMC or NVSM. `sudo nvsm show health`
6. Return the bad fan module using the packaging from the new fan module.

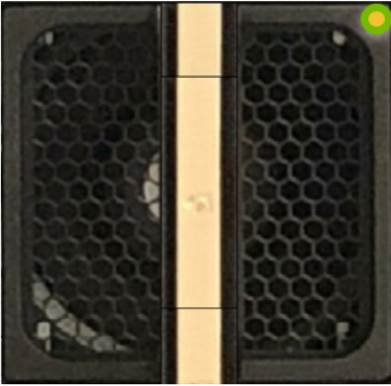
2.2. Identifying the Failed Fan Module

There are several ways to determine the faulty fan module to replace.



2.2.1. Viewing the Fan Module LED

Look for the lit fault LED on the upper right corner of the faulty fan module.



2.2.2. Using the BMC Dashboard and NVSM

1. Identify the faulty fan module using the BMC dashboard.
 1. Log on to the BMC.
 2. Click **Sensor** from the left navigation menu, then review the **Normal Sensors** section.

SPD_FAN_SYS0_F	3654 Rpm
SPD_FAN_SYS0_R	3248 Rpm
SPD_FAN_SYS1_F	3654 Rpm
SPD_FAN_SYS1_R	3248 Rpm
SPD_FAN_SYS2_F	3654 Rpm
SPD_FAN_SYS2_R	3248 Rpm
SPD_FAN_SYS3_F	3654 Rpm
SPD_FAN_SYS3_R	3248 Rpm
SPD_FAN_SYS4_F	4408 Rpm
SPD_FAN_SYS4_R	3886 Rpm
SPD_FAN_SYS5_F	4408 Rpm
SPD_FAN_SYS5_R	3886 Rpm
SPD_FAN_SYS6_F	4408 Rpm
SPD_FAN_SYS6_R	3944 Rpm
SPD_FAN_SYS7_F	4408 Rpm
SPD_FAN_SYS7_R	3886 Rpm

There are two fans in the fan module, identified by SPD_FAN_SYSn_F and SPD_FAN_SYSn_R, where n is the module ID. If either fan fails, then the entire module must be replaced.

2. Use NVSM to confirm the fan issue.

```
$ sudo nvsm show fans
```

In the output, look for the 'unhealthy' status for the same fan.

2.3. Replacing and Returning the Front Fan Module

1. Remove the new fan module from its packaging and be ready to install it.
2. Remove the failed fan module by pressing on the release button on the top of the module and pulling on the handle.



3. Quickly insert the new fan module, observing that the handle release mechanism is facing up.

Caution: Replace the fan module within 30 seconds to prevent overheating of the system components.

4. Confirm that the fan module is healthy working properly by
 - ▶ Verifying that the fan module fault LED is not lit.
 - ▶ Viewing the state of the fan module on the BMC dashboard.
 - ▶ Using NVSM (`sudo nvsm show fans`)
5. Use packaging to pack up the bad fan and follow the shipping instructions to return the bad fan to NVIDIA Enterprise Support.

Chapter 3. Power Supply Replacement

This chapter describes how to replace one of the DGX A100 system power supplies (PSUs).

3.1. Power Supply Replacement Overview

This is a high-level overview of the steps needed to replace a power supply.

1. Identify failed power supply through the BMC and submit a service ticket.
2. Get replacement power supply from NVIDIA Enterprise Support.
3. Identify the power supply using the diagram as a reference and the indicator LEDs.
4. Remove the power cord from the power supply that will be replaced.
5. Replace the failed power supply with the new power supply.
6. Insert the power cord and make sure both LEDs light up green (IN/OUT).
7. Use the BMC to confirm that the power supply is working correctly.
8. Ship back the failed unit to NVIDIA Enterprise Support using the packaging provided.

3.2. Identifying the Failed Power Supply

3.2.1. Identifying the Failed Power Supply from the Back

If physical access to the system is available, you can identify a failed PSU by inspecting the LEDs on the power supply when the system is powered on.



Both LEDs should be solid green. If either of the LEDs are not green or if they are blinking, contact NVIDIA Enterprise Support to troubleshoot the issue.

3.2.2. Identifying the Failed Power Supply from the Console

There are several ways to identify the failed PSU from the DGX A100 console.

- Use the NVSM CLI as follows.

```
$ sudo nvsm show psus
```

The output shows information for each PSU. Look for any that do not report Status_Health=OK.

- View the PSU status from the BMC.

Click **Sensor** from the left side menu and inspect the PSU information from the **Normal Sensors** section.

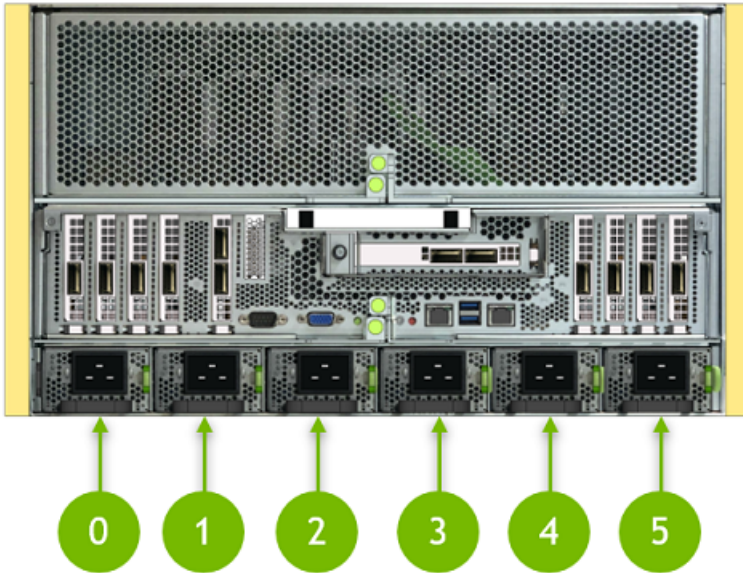
PSU0 STATUS	Presence Detected
PSU1 STATUS	Presence Detected
PSU2 STATUS	Presence Detected
PSU3 STATUS	Presence Detected
PSU4 STATUS	Presence Detected
PSU5 STATUS	Presence Detected
PWR_PSU0	182.000 Watts
PWR_PSU1	182.000 Watts
PWR_PSU2	351.000 Watts
PWR_PSU3	351.000 Watts
PWR_PSU4	260.000 Watts
PWR_PSU5	351.000 Watts
TEMP_PSU0	26.000 °C
TEMP_PSU1	27.000 °C
TEMP_PSU2	30.000 °C
TEMP_PSU3	29.000 °C
TEMP_PSU4	29.000 °C
TEMP_PSU5	27.000 °C

- Use ipmitool.

```
$ sudo ipmitool sdr |grep -i psu
```

Look for power supplies with no temperature reading or an output reading close to or equal to zero.

Both NVSM and the BMC identify each power supply as PSUx, where x is from 0 to 5. The following diagram shows the physical location of each PSU.



3.2.3. Determining the Manufacturer

Important: All PSUs in the system must be from the same manufacturer.

Issue the following to determine the PSU manufacturer.

```
$ sudo nvsm show /chassis/localhost/power/PSUX
```

Where X corresponds to the PSU identifier. The following examples uses PSU0, and shows that the manufacturer is "Delta".

```
$ sudo nvsm show /chassis/localhost/power/PSU0
```

```
/chassis/localhost/power/PSU0
Properties:
FirmwareVersion = 01.05.01.05.01.05
LastPowerOutputWatts = 312
Manufacturer = Delta
MemberId = PSU0
Model = ECD16010092
Name = PSU0
Oem_PSU_Error = <NOT_SET>
PowerSupplyType = AC
SerialNumber = DTHTCP200807M
Status_Health = OK
Status_State = Present
Targets:
Verbs:
cd
show
```

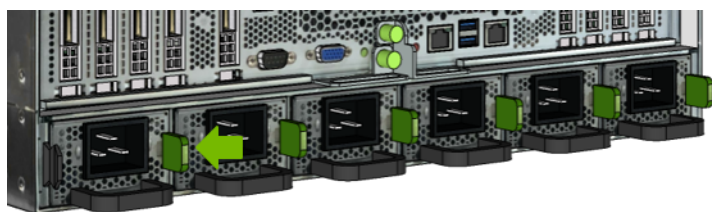
Obtain the replacement PSU (of the same manufacturer) from NVIDIA Enterprise Support.

3.3. Replacing the Power Supply

1. Be sure you have obtained the replacement PSU and that you have saved the packaging to use when sending back the failed PSU.
2. Determine whether you need to shut down the system.
 - ▶ If the three remaining PSUs are working and energized, then you do not need to shut down power to the DGX A100 system.
 - ▶ If fewer than three PSUs are working and energized, then shut down power to the DGX A100 system.
3. Unlock the power cord and then unplug it from the PSU to be replaced.

You may need to dislodge the power cord from the retaining clip.

4. Remove the PSU.
 1. Push on the green tab to release the lock.



2. Pull on the black handle to remove the PSU from the chassis.



5. Install the new power supply.
 1. Insert the new power supply into the chassis and push it all the way in, making sure that the green locking mechanism engages.
 2. Plug in the power cord and lock it in place.
 3. If needed, power on the system.
6. Confirm the installation by
 - ▶ Viewing the PSU status from the BMC dashboard-> ``Sensors`` page.
 - ▶ Running `nvsml show health` to confirm all power supplies are healthy.

Pack the old power supply and ship it back to NVIDIA Enterprise Support.

Chapter 4. Motherboard Tray - Accessing in Place

You will need to access the motherboard tray in order to service the following components. This process provides access to the motherboard components while the motherboard remains attached to the server.

- ▶ M.2 NVMe drives
- ▶ M.2 module riser card
- ▶ Motherboard tray battery
- ▶ Single-port or dual port CX-6 PCI network adapter card

4.1. Accessing the Motherboard Tray

1. Loosen the two motherboard thumbscrews and then pull the handles out to eject the motherboard tray.

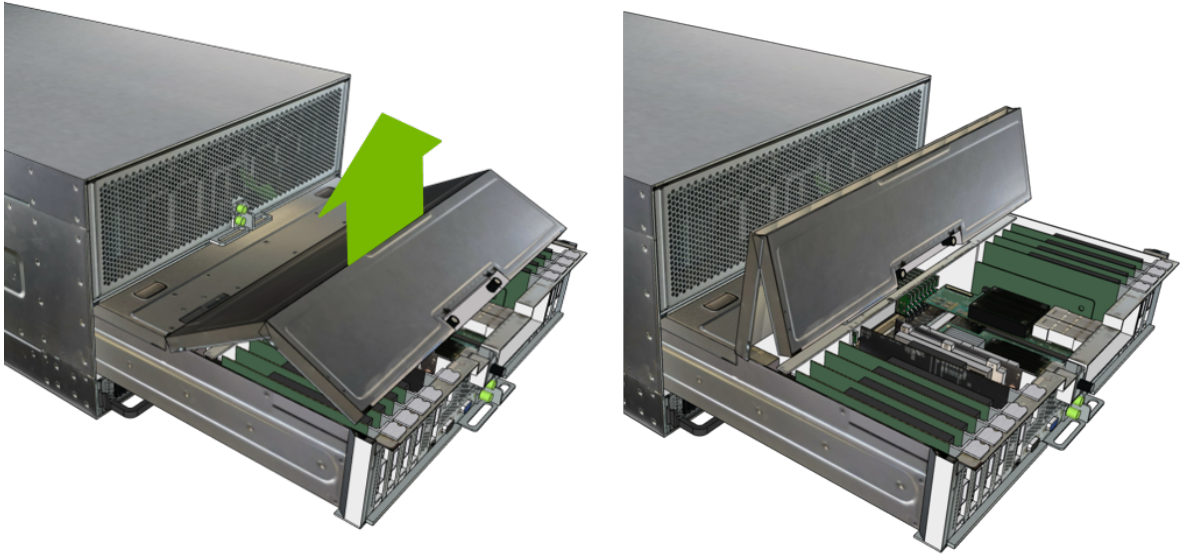




2. Pull the motherboard tray out of the system until it locks, then loosen the two thumbscrews holding the lid in place.

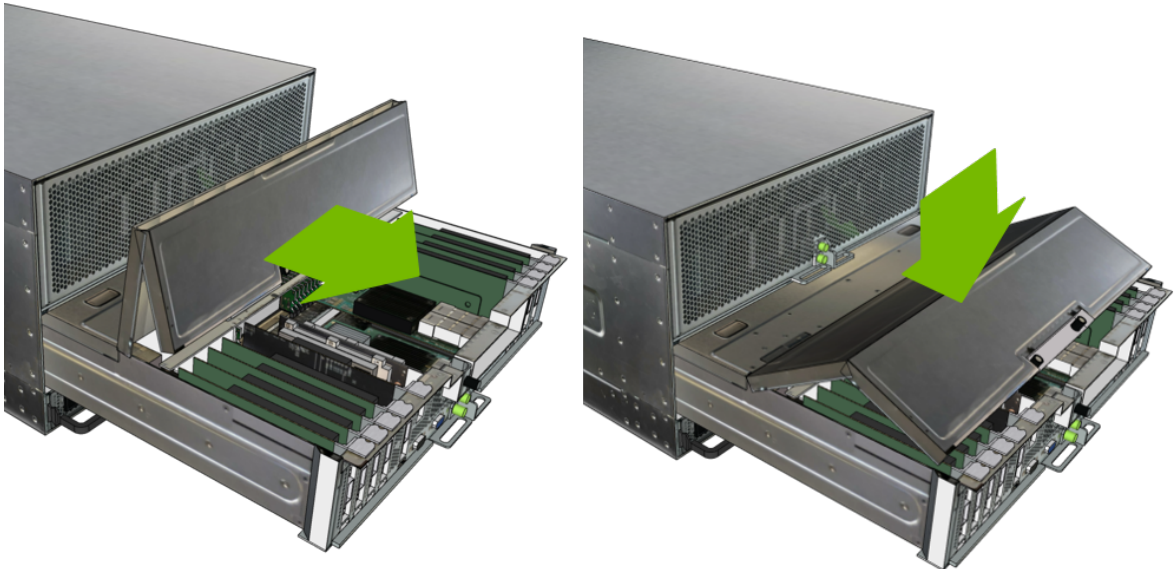


3. Lift the rear section of the motherboard lid.

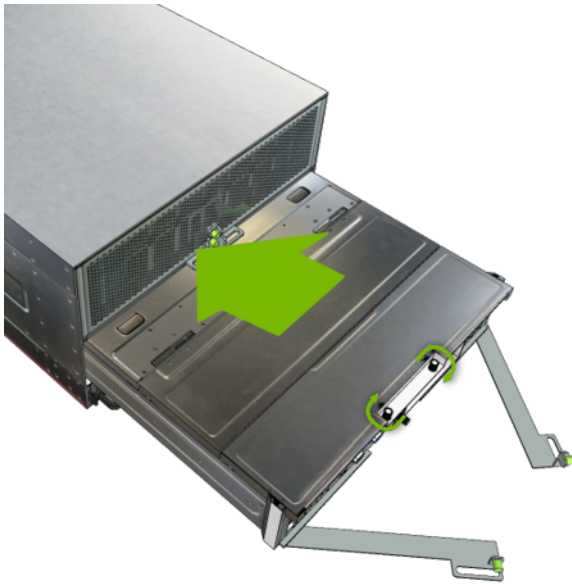


4.2. Replacing the Motherboard Tray

1. Close the lid to the motherboard tray.



2. Tighten the two thumbscrews and then push the motherboard tray into the system.



3. Close the handles to secure the motherboard tray in place.



4. Tighten the motherboard tray thumbscrews to complete the motherboard insertion.



Chapter 5. Motherboard Tray - Removal and Installation

You will need to completely remove the motherboard tray from the server in order to service the following components.

- DIMMs (either adding or replacing)

5.1. Removing the Motherboard Tray

1. Loosen the two motherboard thumbscrews and then pull the handles out to eject the motherboard tray.





2. Pull the motherboard tray out of the system until it locks, then press the two buttons on the top of the lid to release the tray and finish pulling the tray out of the system.



Place the tray on a solid, flat work surface.

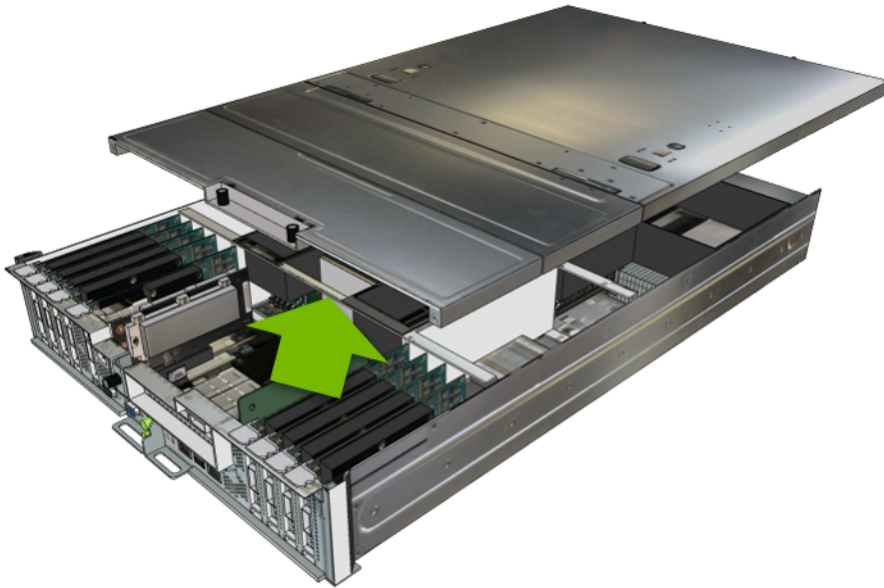
3. Loosen two rear thumbscrews on the motherboard lid.



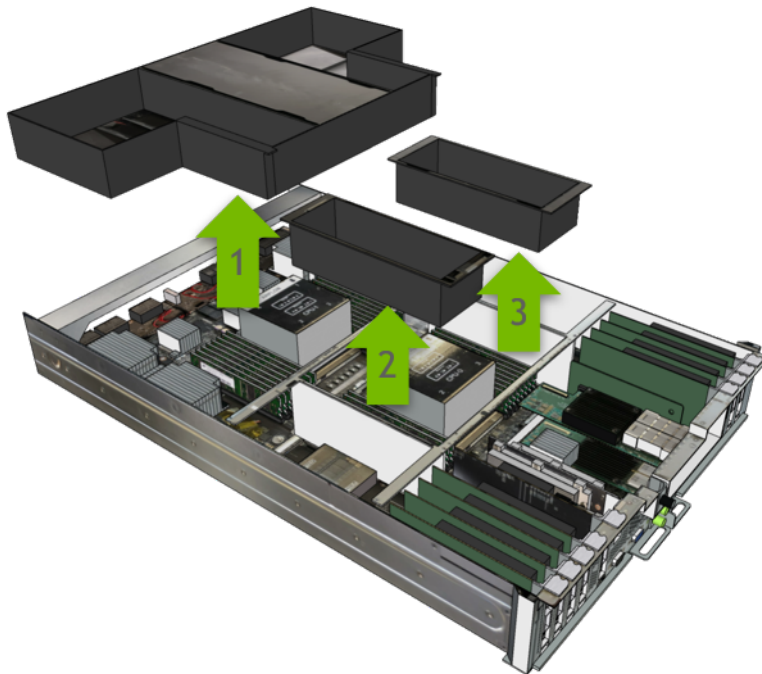
4. Loosen the two front thumbscrews on the motherboard tray lid.



5. Lift the lid off of the tray and set aside.

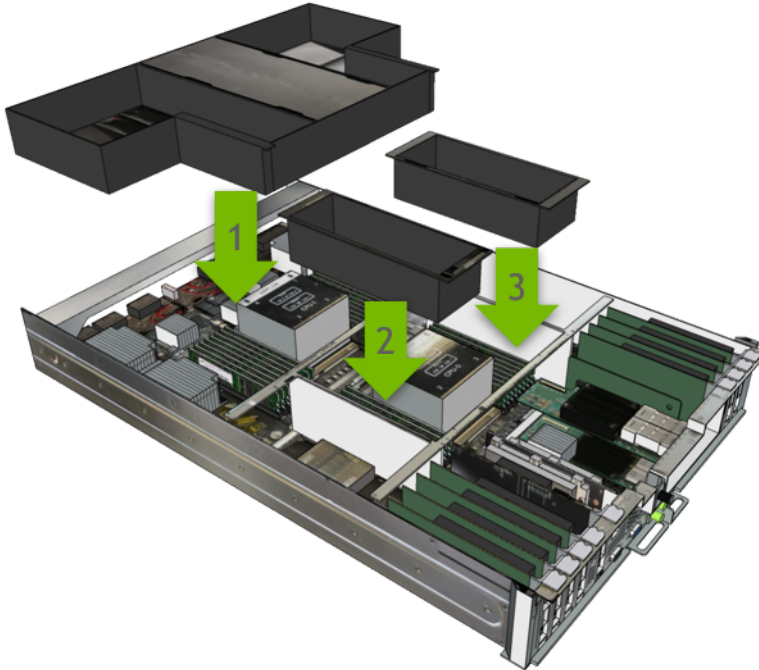


6. Remove all three air baffles to allow access to the DIMMs.



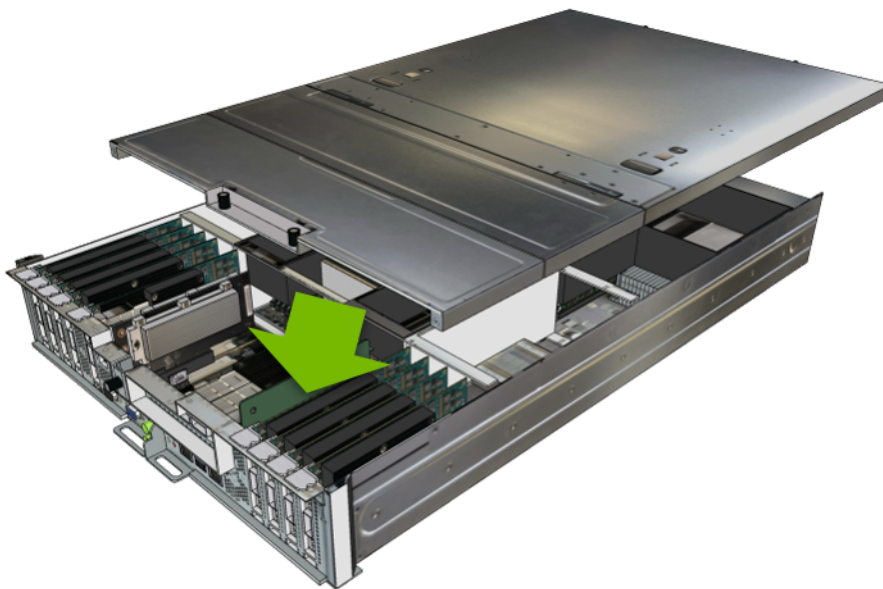
5.2. Reinstalling the Motherboard Tray

1. Reinstall the three air baffles.



2. Replace and secure the lid.

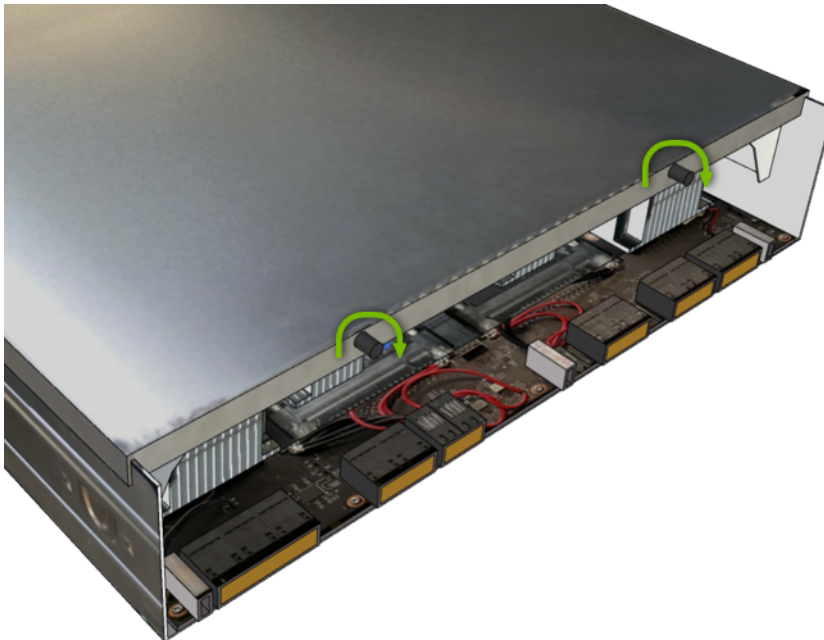
1. Install the lid.



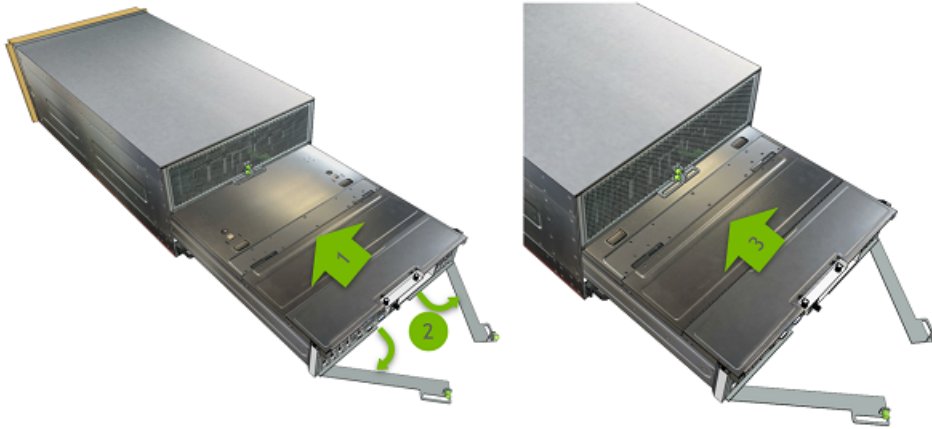
2. Tighten the rear thumbscrews



3. Tighten the front thumbscrews.



3. Slide the motherboard tray into the slot, open the tray handles, and then continue pushing the motherboard tray in.



4. Close the handles to secure the motherboard tray in place.



5. Tighten the motherboard tray thumbscrews to complete the motherboard insertion.



Chapter 6. U.2 NVMe Cache Drive Upgrade from 4 to 8

If you require more drive cache capacity, you can increase the cache by adding four more NVMe drives to the DGX A100 system.

Important: Upgrade using only U.2 NVMe drives of the same density (capacity) as the existing four drives.

6.1. U.2 NVMe Cache Drive Upgrade Overview

This is a high-level overview of the steps needed to upgrade the DGX A100 system's cache size.

1. Identify the density (capacity) of the currently installed NVMe drives.
2. Place an order for additional four NVMe drives from NVIDIA Sales.
3. Power off the system.
4. Remove the air baffles in the blank drive slots.
5. Install the NVMe drives in the DGX A100 system.
6. Power on the system.
7. Re-initialize the `/raid` filesystem to recognize all eight drives.
8. Confirm the system is healthy by running `nvsm show health`.

6.2. Determining the NVMe Drive to Order

1. Identify the drives in the RAID volume by issuing the following.

```
$ sudo nvme list
```

2. Determine the drive size (3.84 TB or 7.68 TB) of the currently installed drives, and order drives of same density (capacity) from the NVIDIA Sales team.

Note: If 3.84 TB drives are installed but you want to use or add 7.68 TB drives, refer to [U.2 NVMe Cache Drive Upgrade to 7.68 TB Drives](#) for instructions.

6.3. Installing the Optional NVMe Drives

1. Be sure you have obtained the additional drives.
2. Back up any critical data to a network shared volume or some other means of backup.
3. Power off the system using the power button or through the console.
4. Remove the blank filler modules from slots 1, 3, 5, and 7.



1. Push the lever release button (on the right side of the lever) to unlock the lever.



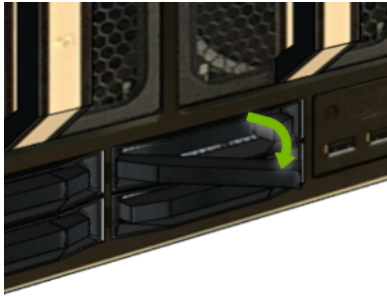
2. Pull the lever to remove the blank filler module.



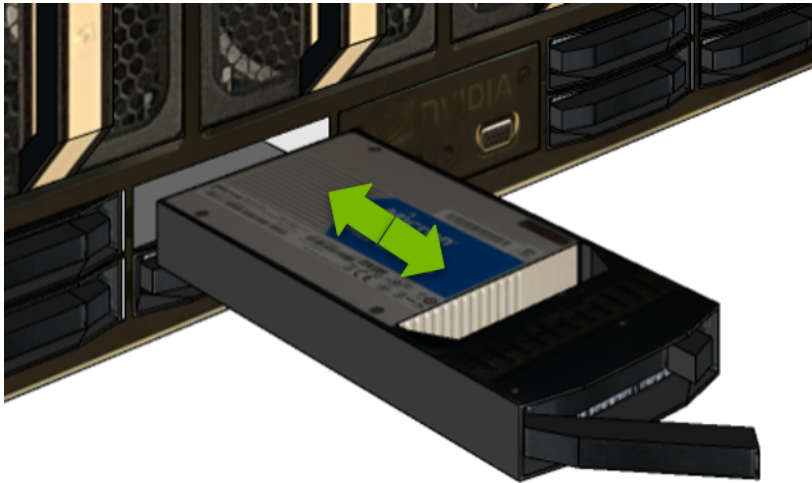
5. Move the drives in slots 4 and 6 and install in slots 1 and 3.



1. Push the lever release button (on the right side of the lever) to unlock the lever.



2. Pull the lever to remove the module.



3. Unlock the release lever and then slide the drive into the slot **until the front face is flush with the other drives.**
4. Close the lever and lock it in place.



6. Install the new NVMe drives in slots 4, 5, 6, and 7.

1. Unlock the release lever and then slide the drive into the slot **until the front face is flush with the other drives.**
2. Close the lever and lock it in place.



7. Power on the system.

Perform the tasks describes in the chapter [U.2 NVMe Cache Drive Post-Installation Tasks](#).

Chapter 7. U.2 NVMe Cache Drive Replacement

Important: Replace only with U.2 NVMe drives of the same manufacturer, model, and density (capacity) as the existing or replaced drives.

7.1. U.2 NVMe Cache Drive Replacement Overview

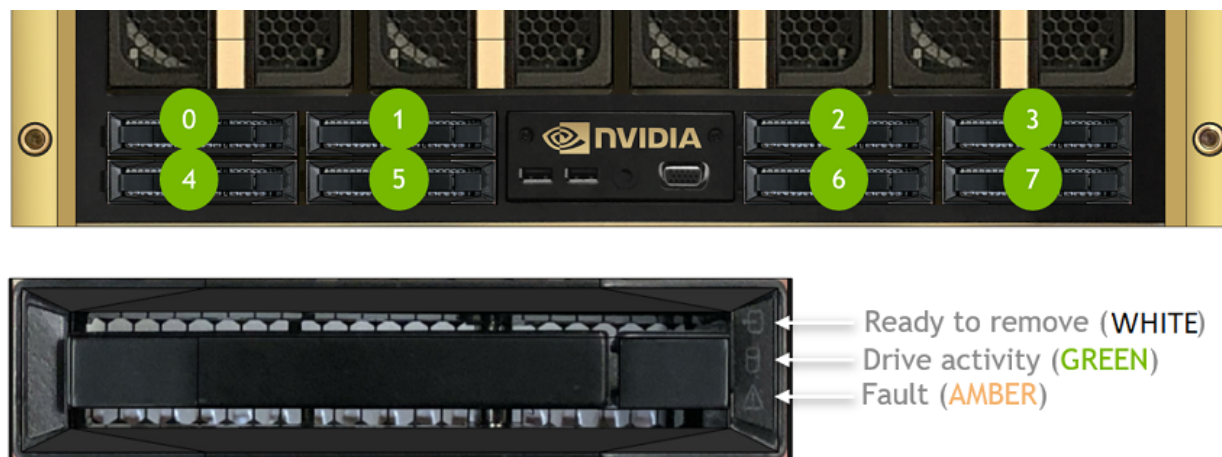
This is a high-level overview of the procedure to replace a cache Non-Volatile Memory Express (NVMe) drive.

1. Identify the failed U.2 NVMe drive.
2. Order a replacement from NVIDIA Enterprise Support.
3. Use `nvsm` to prepare the drive for removal - look for the white LED.
4. Replaced the failed NVMe drive.
5. Rebuild the RAID volume and remount the `/raid` partition.
6. Confirm the system is healthy by running `nvsm show health`.
7. Ship the failed unit back to NVIDIA Enterprise Support using the provided packaging.

7.2. Identifying the Failed U.2 NVMe

7.2.1. Identifying the Failed NVMe from the Front

If physical access to the system is available, you can identify a failed drive by the illuminated amber LED.



7.2.2. Identifying the Failed NVMe from the Console

To identify the failed NVMe drive from the DGX A100 console, enter the following and then look for drive alerts in the output to identify the failed drive.

```
$ sudo nvsm show health
```

The command returns the PCIe bus ID. Refer to the following figure to find the slot ID that corresponds to the PCIe bus ID for the faulty drive.



Fig. 1: NVMe Drives: PCIe to Slot Mapping

Alternatively, you can use the BMC dashboard to access the Sensor screen, the IPMI event log, and the System log to identify issues with the U.2 drives.

Note: The PCIe bus IDs for slots 6 and 7 depend on the firmware version.

7.2.3. Identifying the NVMe Manufacturer and Model

Enter the following, replacing X with the number corresponding to the Linux device name for the failed drive.

```
$ sudo nvsm show /systems/localhost/storage/drives/nvmeXn1
```

Example output:

```
/systems/localhost/storage/drives/nvme5n1
Properties:
  Capacity = 3840755982336
  BlockSizeBytes = 7501476528
  SerialNumber = 174719FCF9F1
  PartNumber = N/A
  Model = Micron_9200_MTFDHAL3T8TCT
  Revision = 100007H0
  Manufacturer = Micron Technology Inc
  Status_State = Enabled
  Status_Health = OK
  Name = Non-Volatile Memory Express
  MediaType = SSD
  IndicatorLED = N/A
  EncryptionStatus = N/A
  HotSpareType = N/A
  Protocol = NVMe
  NegotiatedSpeedsGbs = 0
  Id = 5
```

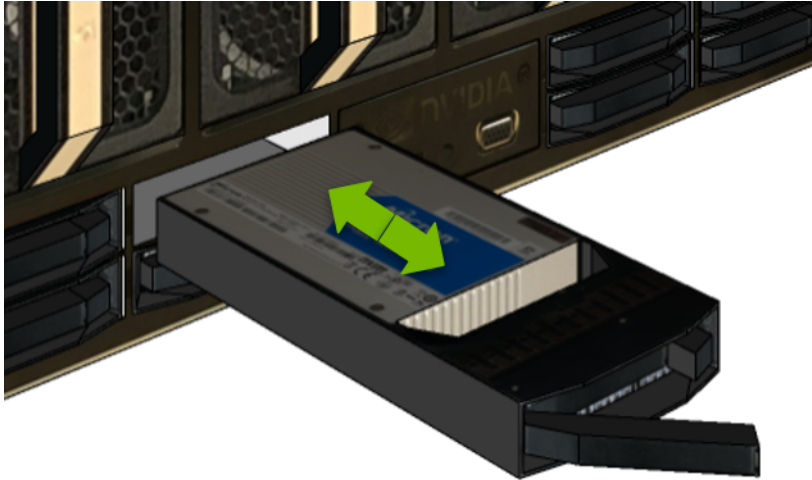
Determine the manufacturer and model from the 'Model' entry in the output, and then request a replacement NVMe from NVIDIA Enterprise Support, specifying this information.

7.3. Replacing the U.2 NVMe Drive

1. Be sure you have requested and obtained the replacement drive from NVIDIA Enterprise Support.
2. Back up any critical data to a network shared volume or some other means of backup.
3. Power off the system using the power button.
4. Remove the NVMe drive.
 1. Push the lever release button (on the right side of the lever) to unlock the lever.



2. Pull the lever to remove the module.



5. Replace the new NVMe drive in the same slot.
 1. Unlock the release lever and then slide the drive into the slot **until the front face is flush with the other drives.**
 2. Close the lever and lock it in place.



6. Power on the system.

Perform the tasks describes in the chapter [U.2 NVMe Cache Drive Post-Installation Tasks](#).

Chapter 8. U.2 NVMe Cache Drive Upgrade to 7.68 TB Drives

If you require more drive cache capacity, you can increase the cache by replacing the existing 3.84 TB drives with 7.68 TB drives.

8.1. U.2 NVMe Cache Drive 7.68 TB Upgrade Overview

This is a high-level overview of the steps needed to upgrade the DGX A100 system's cache size.

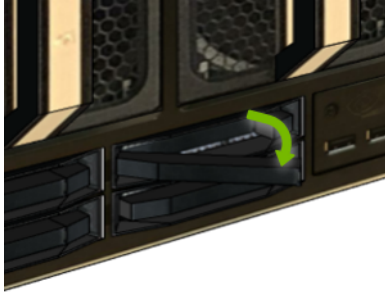
1. Place an order for the 7.68 TB U.2 NVMe drives from NVIDIA Sales.
2. Power off the system.
3. Remove the existing components.
 - ▶ Remove all 3.84 TB cache drives.
 - ▶ If you are also upgrading from four to eight cache drives, then remove the air baffles in the blank drive slots.
4. Install the 7.68 TB NVMe drives in the DGX A100 system.
5. Power on the system.
6. Re-initialize the `/raid` filesystem to recognize all eight drives.
7. Confirm the system is healthy by running `nvsm show health`.

8.2. Installing the 7.68 TB NVMe Drives

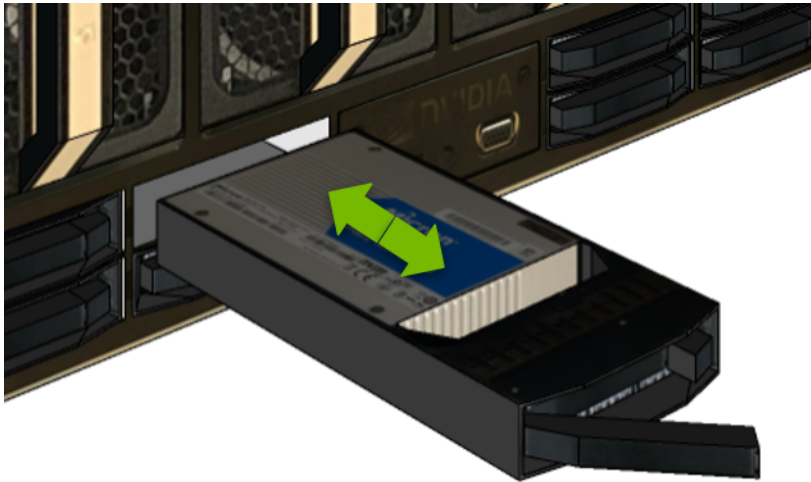
1. Be sure you have obtained the additional drives.
2. Back up any critical data to a network shared volume or some other means of backup.
3. Power off the system using the power button or through the console.
4. Remove all 3.84 TB NVMe drives.

Important: You must remove all 3.84 TB drives so that there is no mix of 3.84 TB and 7.68 TB drives in the same system.

1. Push the lever release button (on the right side of the lever) to unlock the lever.



2. Pull the lever to remove the module.



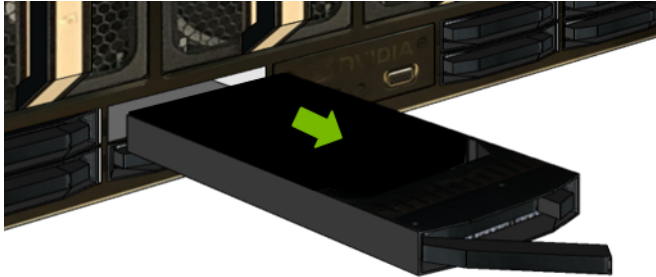
5. (Optional) If you are also upgrading from four to eight drives, remove the blank filler modules from slots 1, 3, 5, and 7.



1. Push the lever release button (on the right side of the lever) to unlock the lever.



2. Pull the lever to remove the blank filler module.



6. Install the 7.68 TB NVMe drives.

1. Unlock the release lever and then slide the drive into the slot **until the front face is flush with the other drives.**
2. Close the lever and lock it in place.



7. Power on the system.

Perform the tasks describes in the chapter [U.2 NVMe Cache Drive Post-Installation Tasks](#).

Chapter 9. U.2 NVMe Cache Drive Post-Installation Tasks

This chapter describes the tasks that are typically needed after replacing a U.2 NVME drive or upgrading from 4 to 8 drives.

9.1. Recreating the Cache RAID 0 Volume

1. Power on the system and log in.
2. Confirm that all expected drives are visible.

```
$ sudo nvme list
```

Two boot drives and either four or eight cache drives should be visible, depending on how many are installed in the system.

3. If the previously installed cache drives (which are capable of self-encryption) are locked with an access key, then unlock them.

```
$ sudo nv-disk-encrypt disable
```

4. Re-create the cache volume and the `/raid` filesystem.

```
$ sudo nvsm start /systems/localhost/storage/volumes/rebuild
```

Output (with appropriate responses to prompts)

```
PROMPT: In order to rebuild volume, volume type is required.
Please specify the volume type to rebuild.
(Options: raid-0 only) Type of volume rebuild (CTRL-C to cancel): raid-0

WARNING: Once the RAID-0 rebuild process is started,
all data currently stored on raid will be lost. Start RAID-0 rebuild on raid? [y/
↵n] y
/systems/localhost/storage/volumes/rebuild started at 2021-01-27 16:03:41.093694
Finished rebuilding RAID-0 on volume raid
100.0% [=====]
Status: Done
```

5. If the volume is to be locked with an access key, then re-enable drive encryption.
 - To enter drive passwords yourself, issue the following.

```
$ sudo nv-disk-encrypt init -g
```

The software prompts you to enter a password for the vault, and then a password for each eligible SED.

Passwords must consist of only upper-case letters, lower-case letters, digits, and/or the following special-characters: ~ : @ % ^ + = _ ,

```
$ sudo nv-disk-encrypt lock
```

- To allow the encryption software to randomly generated the passwords, issue the following.

```
$ sudo nv-disk-encrypt init -k <your-vault-password> -g -r
```

The vault password must consist of only upper-case letters, lower-case letters, digits, and/or the following special-characters: ~ : @ % ^ + = _ ,

```
$ sudo nv-disk-encrypt lock
```

```
$ sudo nvsm enable /systems/localhost/storage/volumes/md1/encryption
```

6. Issue the following to confirm the volume is healthy and that the system is healthy.

```
$ sudo nvsm show storage
```

```
$ sudo nvsm show health
```

Make sure that the drive firmware is up to date. Refer to the [DGX A100 firmware release notes](#) for information on the latest firmware for the U.2 NVMe drive.

9.2. Returning the NVMe Drive

Use the packaging from the new drive and follow the instructions that came with the package to ship the old drive back to NVIDIA Enterprise Support.

Note: If your organization has purchased a media retention policy, you may be able to keep failed drives for destruction. Check with NVIDIA Enterprise Support on the status of the policy for specifics.

Chapter 10. M.2 NVMe Boot Drive Replacement

10.1. M.2 NVMe Boot Drive Replacement Overview

This is a high-level overview of the procedure to replace a boot drive.

1. With the help of NVIDIA Enterprise Support, determine which M.2 drive needs to be replaced.
2. Get replacement from NVIDIA Enterprise Support.
3. Power down the system.
4. Label all cables and unplug them from the motherboard tray.
5. Slide motherboard out until it locks in place.
6. Open rear compartment and pull out the M.2 riser card with both M.2 disks attached.
7. Replace the failed M.2 device on the riser card.
8. Install the M.2 riser card with both M.2 disks.
9. Close the rear motherboard compartment and then slide the motherboard back into the system.
10. Plug in all cables using the labels as a reference.
11. Power on the system.
12. Confirm the M.2 RAID 1 mirror is synchronizing.
13. Ship back the failed unit to NVIDIA Enterprise Support using the packaging provided.

10.2. Identifying the Failed M.2 NVMe

The DGX A100 system automatically sets the failed M.2 drive offline when it detects the failure.

1. Identify which of the M.2 drives has failed (nvme0n1 or nvme1n1).

```
$ sudo nvsm show health
```

2. You can confirm this by issuing the following.

```
$ sudo mdadm -D /dev/md0
```

Normally, the output would show both drives (nvme0 and nvme1) in an active sync state. The following example output shows only nvme1 in active sync, indicating that nvme0n1 is the failed drive.

Number	Major	Minor	RaidDevice	State	
0	259	2	0	active sync	/dev/nvme1n1p2
-	0	0	1	removed	

3. Make a note of the device name for the failed drive (nvme0 or nvme1) and the device name for the good drive (nvme0 or nvme1).

You will need this information when rebuilding the RAID 1 array after replacing the drive.

4. Obtain the replacement from NVIDIA Enterprise Support.

10.3. Replacing the M.2 NVMe Drive

Before attempting to replace one of the M.2 NVMe drives, be sure to have performed the following:

- Determined the location ID of the faulty M.2 NVMe drive.
- Obtained the replacement M.2 NVMe drive and have saved the packaging for use when returning the faulty drive.

M.2 NVMe Drives:

► **40GB model**

- PCIe Bus: 22 -> /dev/nvme1
- PCIe Bus: 23 -> /dev/nvme2

► **80GB model**

- PCIe Bus: 22 -> /dev/nvme2
- PCIe Bus: 23 -> /dev/nvme3

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

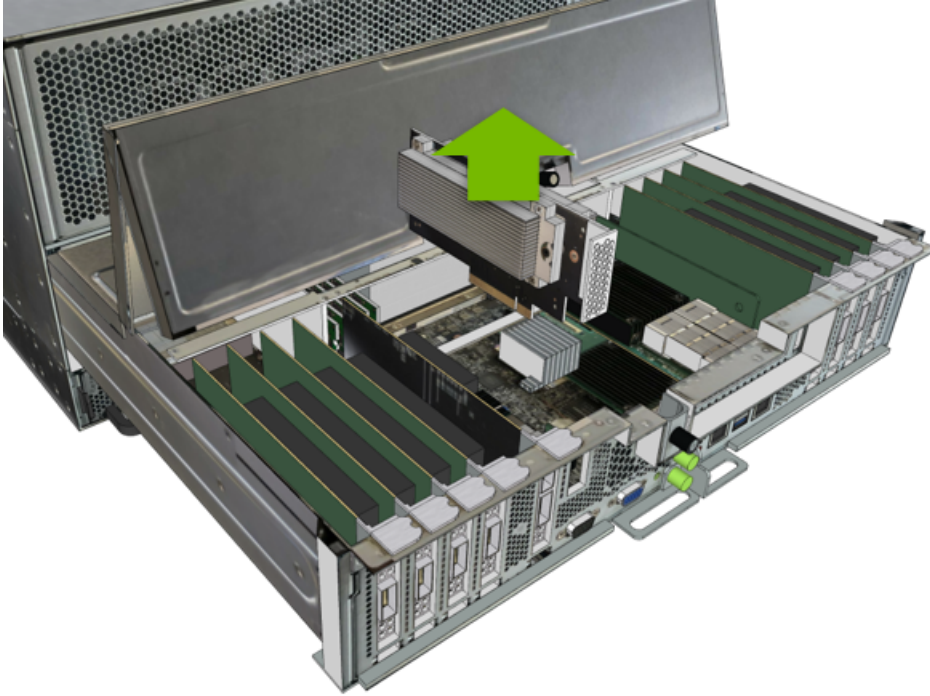
1. Back up any critical data to a network shared volume or some other means of backup.
2. If not already done, mark the drive as failed, then remove the failed drive from the array by issuing the following (replacing X with the failed drive identifier - 0 or 1).

```
$ sudo mdadm --manage /dev/md0 --fail /dev/nvme<X>n1
```

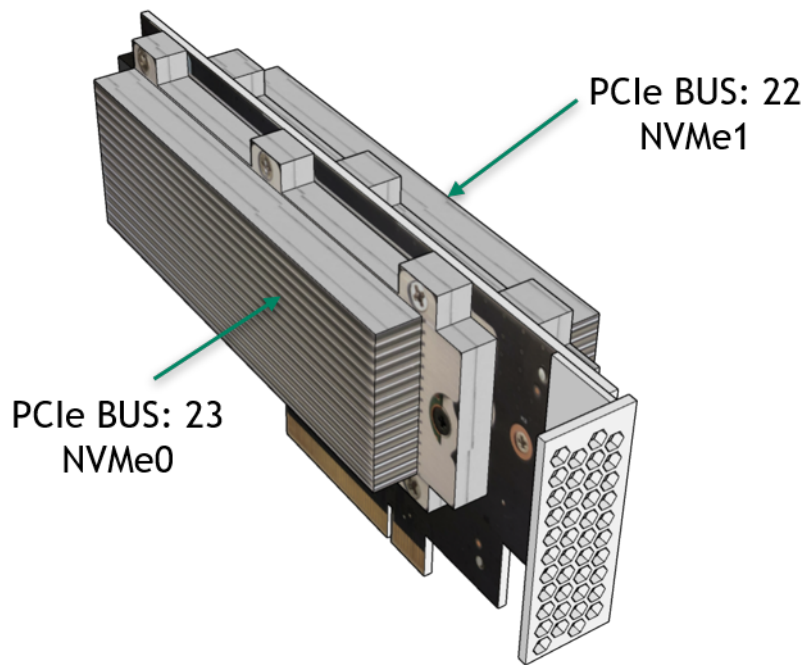
```
$ sudo mdadm --manage /dev/md0 --remove /dev/nvme<X>n1
```

3. Power down the system.

4. Label all network, monitor, and USB cables connected to the motherboard tray for easy identification when reconnecting.
5. Unplug all power cords, and all network, monitor, and USB cables.
6. Remove the motherboard tray.
Refer to the instructions in the section [Accessing the Motherboard Tray](#).
7. Remove the M.2 riser card from the motherboard tray by lifting the riser assembly.



8. Identify the failed M.2 module and remove it from the riser card.
 1. Determine the location of the drive to remove.

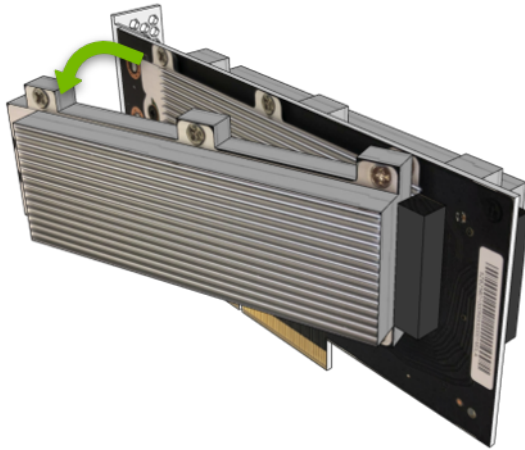


2. Using a Phillips #1 screwdriver, loosen the black screw that secures the drive in place.

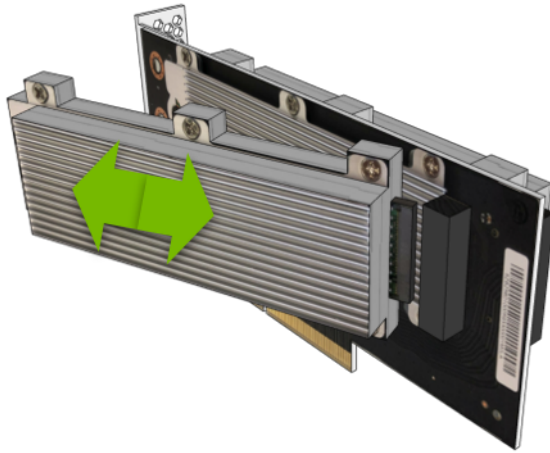
Note: The screw is not a captive screw and can drop. Be careful when loosening the screw to avoid dropping and losing the screw.



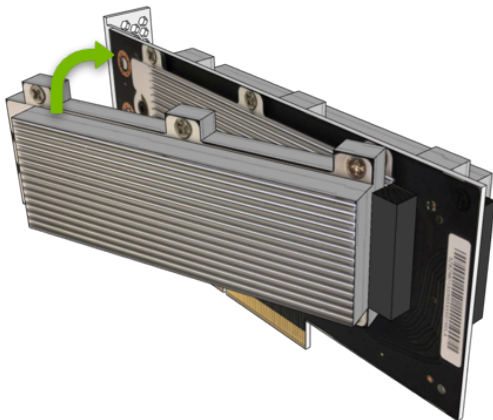
3. Carefully lift the M.2 drive away from the riser.



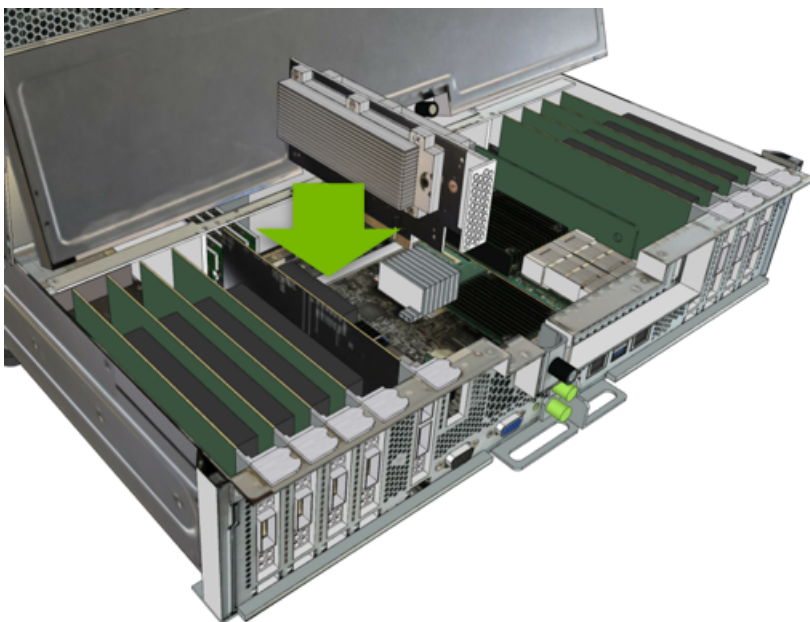
4. Pull the drive to disconnect from the connector on the riser board, then insert the new drive into the connector on the riser board.



5. Place the drive against the card and secure by tightening the screw using a Phillips #1 screwdriver.



9. Install the assembled module on the motherboard by inserting the riser card in its slot.



10. Close the motherboard tray lid and then install the motherboard tray.

Refer to the instructions in the section [Replacing the Motherboard Tray](#).

11. Connect all the cables to the motherboard tray, connect all power cords, then power on the system.

Rebuild the RAID 1 array according to the instruction in the section [Rebuilding the Boot Drive RAID 1 Volume](#).

10.4. Rebuilding the Boot Drive RAID 1 Volume

After replacing a faulty M.2 OS drive, you must rebuild the RAID 1 array.

1. If you have not already done so, boot the DGX A100 system and log in.
2. Rebuild the boot drive mirror.

In the following steps, replace X with the number that corresponds to the replaced drive. If you did not note this information when identifying the failed drive, then follow the instructions in the first step of [Identifying the Failed M.2 NVMe](#).

1. Start the rebuild process.

```
$ sudo nvsm start /systems/localhost/storage/volumes/md0/rebuild/
```

2. Enter the device name of the spare (replaced) drive when prompted - either nvme0n1 or nvme1n1, depending on which drive was replaced.

```
PROMPT: In order to rebuild this volume, a spare drive
        is required. Please specify the spare drive to
        use to rebuild md0.
Name of spare drive for md0 rebuild (CTRL-C to cancel): nvmeXn1
WARNING: Once the volume rebuild process is started, the
```

(continues on next page)

(continued from previous page)

```
process cannot be stopped.  
Start RAID-1 rebuild on md0? [y/n] y
```

After entering **y** at the prompt to start the RAID 1 rebuild, the “Initiating rebuild ...” message appears.

```
/systems/localhost/storage/volumes/md0/rebuild started at 2018-10-12  
15:27:26.525187  
Initiating RAID-1 rebuild on volume md0...  
  0.0% [ \                               ]
```

After about 30 seconds, the “Rebuilding RAID-1 ...” message should appear.

```
/systems/localhost/storage/volumes/md0/rebuild started at 2018-10-12  
15:27:26.525187  
Rebuilding RAID-1 rebuild on volume md0...  
 31.0% [ =====/                        ]
```

If this message remains at “Initiating RAID-1 rebuild” for more than 30 seconds, then there is a problem with the rebuild process. In this case, make sure the name of the replacement drive is correct and try again.

The RAID 1 rebuild process should take about 1 hour to complete.

Make sure that the drive firmware is up to date. Refer to the [DGX A100 firmware release notes](#) for information on the latest firmware for the M.2 NVMe drive.

10.5. Returning the NVMe Drive

Use the packaging from the new drive and follow the instructions that came with the package to ship the old drive back to NVIDIA Enterprise Support.

Note: If your organization has purchased a media retention policy, you may be able to keep failed drives for destruction. Check with NVIDIA Enterprise Support on the status of the policy for specifics.

Chapter 11. M.2 Boot Drive Riser Assembly Replacement

This chapter applies when both M.2 OS drives need to be replaced. In that case, a replacement riser assembly (which includes both M.2 NVMe drives) should be ordered.

11.1. M.2 Boot Drive Riser Assembly Replacement Overview

This is a high-level overview of the procedure to replace the boot drive riser assembly.

1. Confirm that both M.2 drives cannot be reached and need to be replaced.
2. Get replacement M.2 riser assembly from NVIDIA Enterprise Support.
3. Power down the system.
4. Label all motherboard tray cables and unplug them.
5. Slide out the motherboard tray and open the motherboard tray lid.
6. Pull out the M.2 riser assembly with both M.2 disks attached.
7. Install the new M.2 riser assembly with both M.2 disks attached..
8. Close the lid on the motherboard tray.
9. Slide the motherboard tray into the system.
10. Plug in all cables using the labels as a reference.
11. Power on the system.
12. Re-install the OS and confirm the system is healthy.
13. Ship back the failed unit to NVIDIA Enterprise Support using the packaging provided.

11.2. Determining a Failed M.2 NVMe Riser Assembly

The following are the conditions for which NVIDIA Enterprise Support may instruct the M.2 riser assembly be replaced:

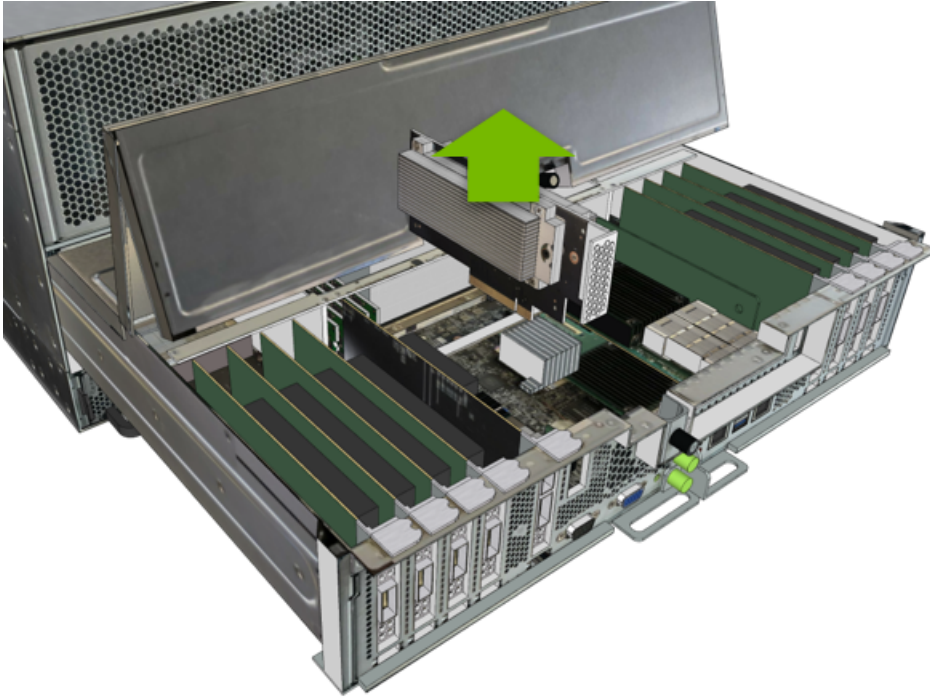
- ▶ The DGX A100 cannot be booted.
- ▶ The boot drives cannot be seen from the SBIOS.
- ▶ The system indicates that the boot drives are not available when booting from the ISO image.
- ▶ Loss of communication with the M.2 boot drives.
- ▶ The M.2 riser assembly was damaged.

11.3. Replacing the M.2 NVMe Riser Assembly

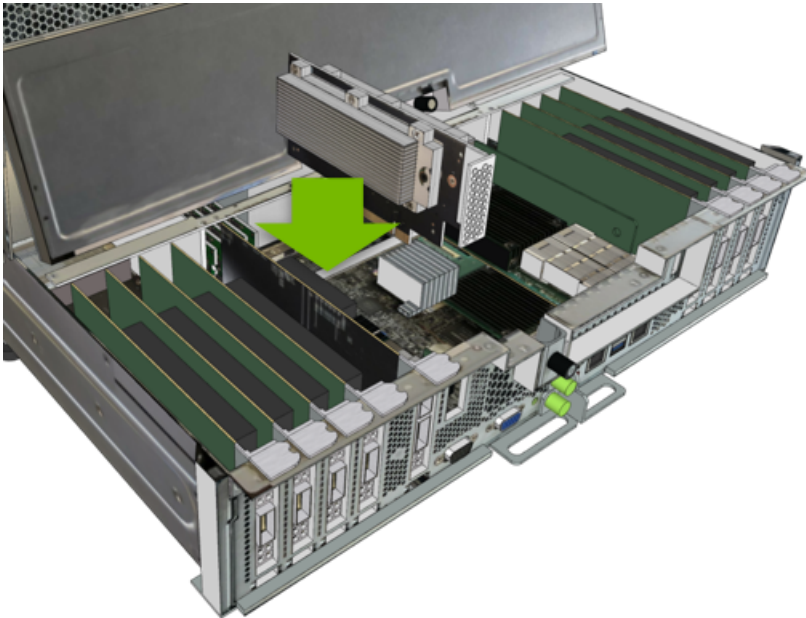
Before attempting to replace the M.2 NVMe riser assembly, be sure you have obtained the replacement assembly and have saved the packaging for use when returning the faulty riser assembly.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

1. Power down the system.
You will likely need to use the BMC console.
2. Label all cables connected to the motherboard tray for easy identification when reconnecting.
3. Remove the motherboard tray.
Refer to the instructions in the section [Accessing the Motherboard Tray](#).
4. Remove the M.2 riser card from the motherboard tray by lifting the riser assembly.



5. Install the assembled module on the motherboard by inserting the riser card in its slot.



6. Close the motherboard tray lid and then install the motherboard tray.
Refer to the instructions in the section [Replacing the Motherboard Tray](#).
7. Connect all the cables to the motherboard tray.
8. Re-install the DGX OS server software.
See the [DGX A100 User Guide](#) for detailed instructions.

11.4. Returning the Riser Assembly

Use the packaging from the new riser assembly and follow the instructions that came with the package to ship the old riser assembly back to NVIDIA Enterprise Support.

Note: If your organization has purchased a media retention policy, you may be able to keep failed drives for destruction. Check with NVIDIA Enterprise Support on the status of the policy for specifics.

Chapter 12. DIMM Replacement

12.1. DIMM Replacement Overview

This is a high-level overview of the procedure to replace a dual inline memory module (DIMM) on the DGX A100 system.

1. Use the `nvsm health` commands to identify the failed DIMM
2. Get a replacement DIMM from NVIDIA Enterprise Support.
3. Shut down the system.
4. Label all motherboard tray cables and unplug them.
5. Remove the motherboard tray and place on a solid flat surface.
6. Remove the motherboard tray lid.
7. Use the reference diagram on the lid of the motherboard tray to identify the failed DIMM.
8. Replace the bad DIMM with the new one.
9. Close the lid on the motherboard tray.
10. Insert the motherboard tray into the system.
11. Plug in all cables using the labels as a reference.
12. Power on the system.
13. Verify that all DIMMs are now healthy with `nvsm`.

12.2. Identifying the Failed DIMM

1. From the console, run the following `nvsm` command to identify memory alerts.

```
$ sudo nvsm show /systems/localhost/memory/alerts
```

Alerts will appear under the Target section. For example.

```
Targets:
    alert0
```

2. Get specific information about the memory alert.
The following example obtains information for `alert0`.

```
$ sudo nvsm show /systems/localhost/memory/alerts/alert0
```

Inspect the `component_id` = line to determine the DIMM ID. The following example shows a DIMM ID of A1.

```
Properties:
system_name = ....
component_id = CPU1_DIMM_A1
...
```

The output provides other information about the alert that can be provided to NVIDIA Enterprise Support.

3. Determine the DIMM manufacturer.

```
$ sudo nvsm show memory
```

4. Request the replacement DIMM from NVIDIA Enterprise Support, specifying the manufacturer.

12.3. Replacing the DIMM

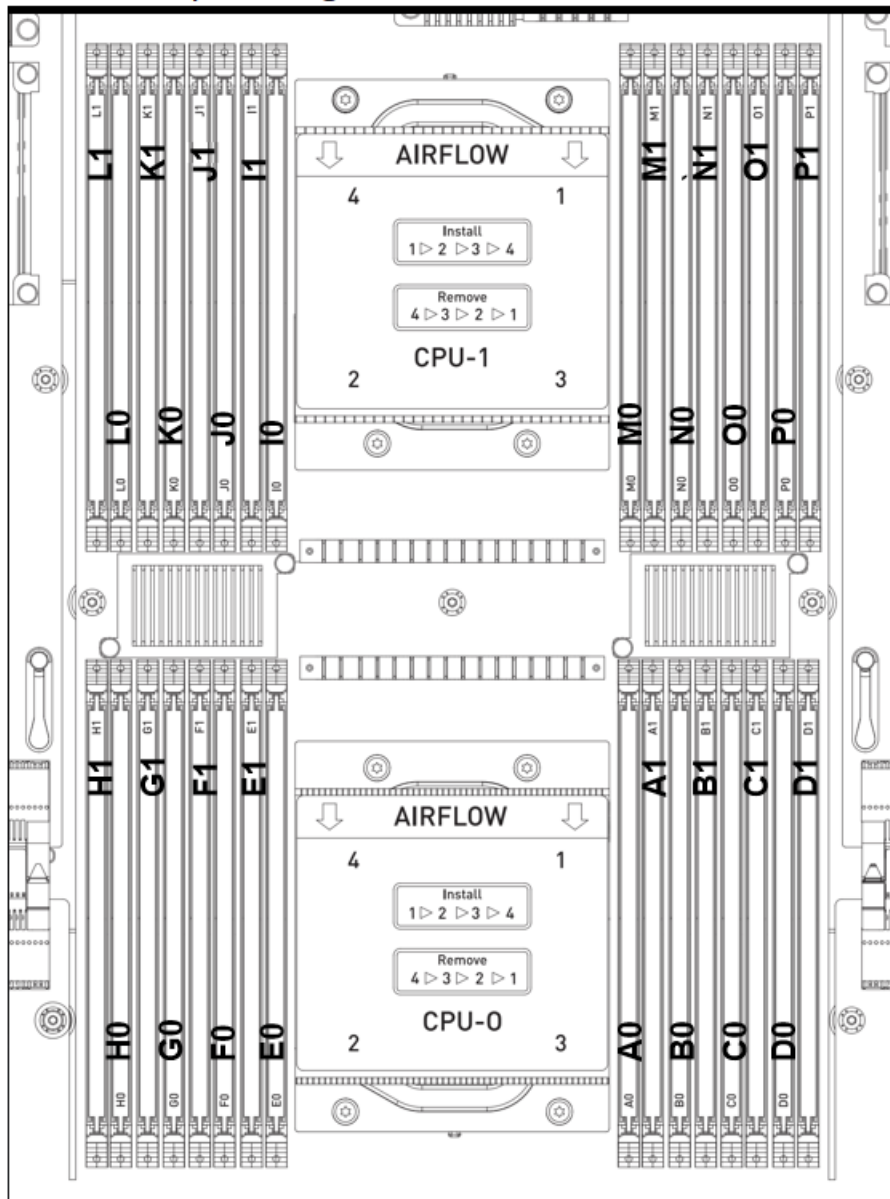
Before attempting to replace any of the dual inline memory modules (DIMMs), be sure to have performed the following:

- Determined the location ID of the faulty DIMM needing replacement as explained in: *ref:identifying-failed-dimm*. The location ID is an alphanumeric designator, such as A0, A1, B0, B1, etc.
- Obtained the replacement DIMM and have saved the packaging for use when returning the faulty DIMM.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

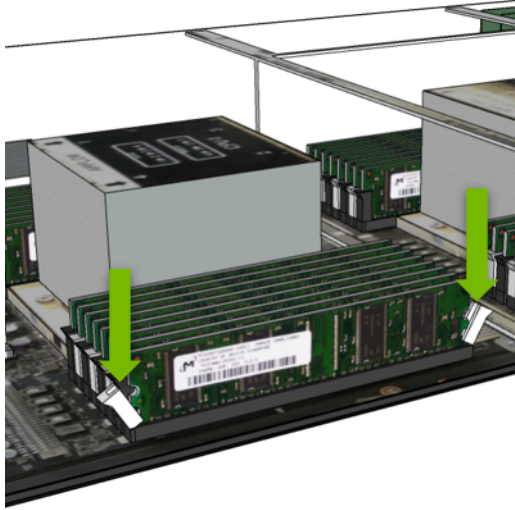
1. Power down the system.
2. Label all cables connected to the motherboard tray for easy identification when reconnecting.
3. Remove the motherboard tray.
Refer to the instructions in the section *Removing the Motherboard Tray*.
4. Using the diagram label on the lid as a guide, locate the faulty DIMM to be replaced.

Top of diagram = front of server

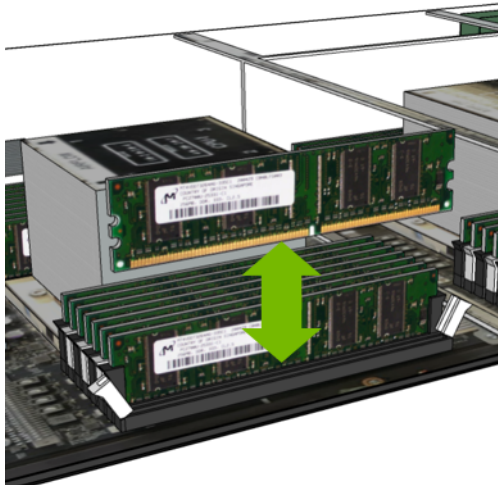


5. Remove the DIMM.

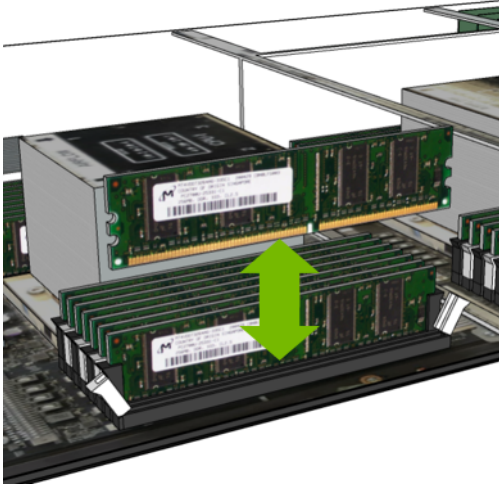
1. Press down on the side latches at both ends of the DIMM socket to push them away from the DIMM. This should unseat the DIMM from the socket.



2. Pull the DIMM straight up to remove it from the socket.



6. Carefully insert the replacement DIMM.
 1. Make sure the socket latches are open.
 2. Position the DIMM over the socket, making sure that the notch on the DIMM lines up with the key in the slot, then press the DIMM down into the socket until the side latches click in place.
 3. Make sure that the latches are up and locked in place.



7. Install the three air baffles, replace the motherboard tray lid and then install the motherboard tray.

Refer to the instructions in the section [Reinstalling the Motherboard Tray](#).

8. Connect all the cables to the motherboard tray.
9. Install all the power cords.
10. Power on the system and log in.
11. Confirm that the system is healthy.

```
$ sudo nvsm show health  
$ sudo nvsm show /systems/localhost/memory/alerts
```

There should be no new alerts listed.

12. Ship the bad DIMM back to NVIDIA Enterprise Support.

Chapter 13. DIMM Upgrade

You can upgrade the system memory by installing 16 additional DIMMs.

13.1. DIMM Upgrade Overview

This is a high-level overview of the procedure to add 16 additional dual-inline memory modules (DIMMs) on the DGX A100 system.

1. Obtain the memory upgrade (16 DIMMs) from NVIDIA Sales.
2. Shut down the system.
3. Label all motherboard tray cables and unplug them.
4. Remove the motherboard tray and place on a solid flat surface.
5. Remove the motherboard tray lid.
6. Use the reference diagram on the lid of the motherboard tray to identify the empty DIMM locations (spots with air baffles installed).
7. Replace the air baffles with the new DIMMs.
8. Close the lid on the motherboard tray.
9. Insert the motherboard tray into the system.
10. Plug in all cables using the labels as a reference.
11. Power on the system.
12. Verify that all DIMMs as well as the system are healthy using `nvsm`.

13.2. Upgrading the DIMM

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

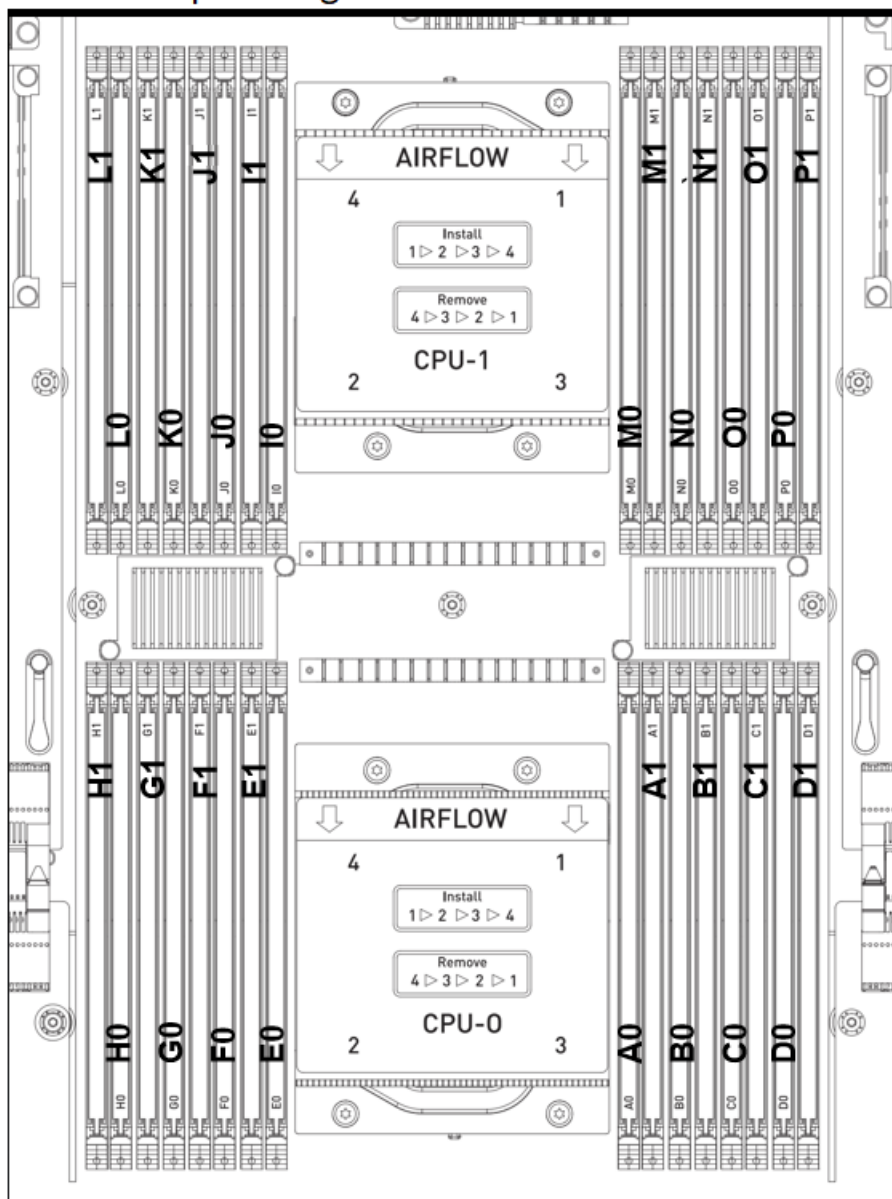
1. Power down the system.

2. Label all cables connected to the motherboard tray for easy identification when reconnecting.
3. Remove the motherboard tray and air baffles.

Refer to the instructions in the section *Removing the Motherboard Tray*.

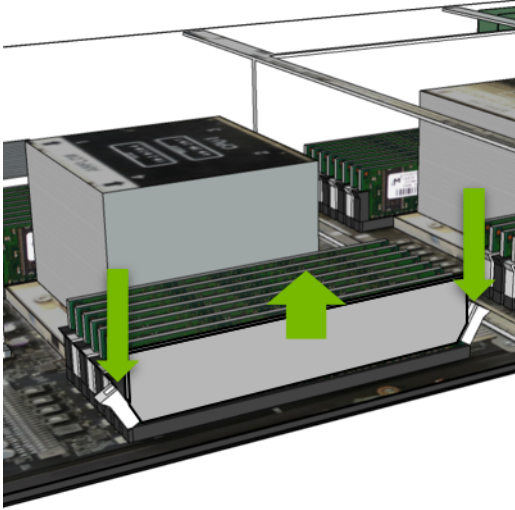
4. Using the diagram label on the lid as a guide, locate the DIMMs to be installed during the upgrade.

Top of diagram = front of server



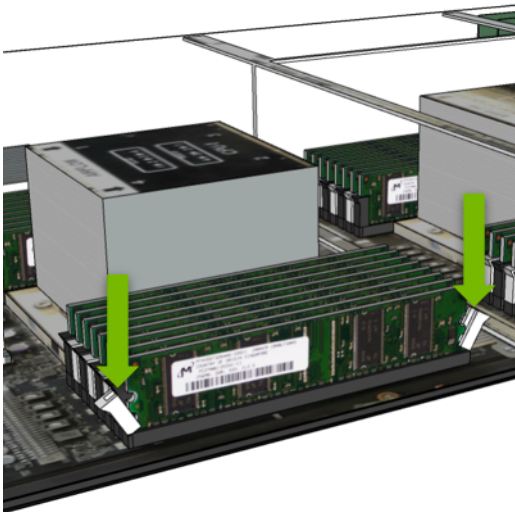
5. Remove the DIMM air baffles.

Press down on the side latches at both ends of the air baffle to eject the module from the slot, then pull the air baffle out of the slot.

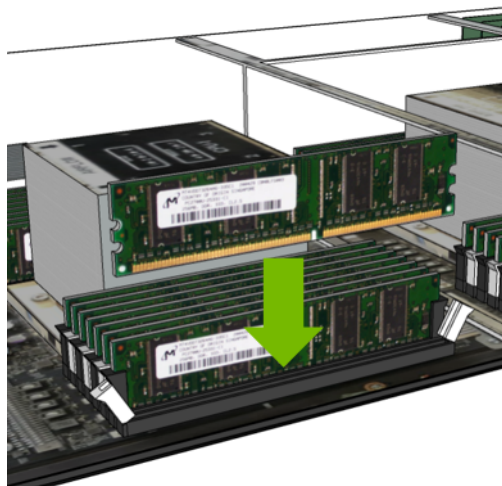


6. Remove 8 DIMMs from CPU-1 slots I1, J1, K1, L1, M, N1, O1, and P1

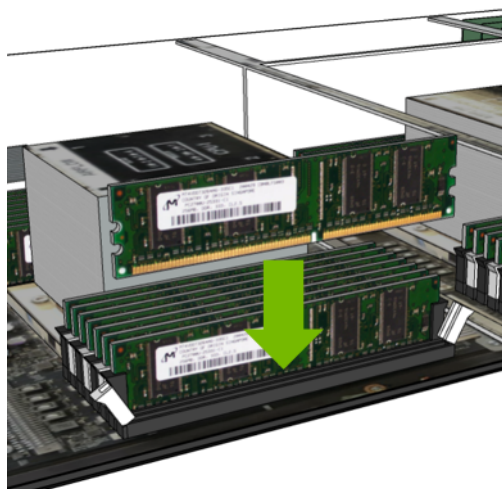
Press down on the side latches at both ends of the DIMM to eject the module from the slot, then pull the DIMM out of the slot.



7. Carefully insert the 8 DIMMs that you just removed into CPU-0 slots A0, B0, C0, D0, E0, F0, G0, and H0.
1. Make sure the socket latches are open.
 2. Position the DIMM over the socket, making sure that the notch on the DIMM lines up with the key in the slot, then press the DIMM down into the socket until the side latches click in place.



3. Make sure that the latches are up and locked in place.
8. Install the new DIMMs from the upgrade kit to CPU-1 slots I0, I1, J0, J1, K0, K1, L0, L1, M0, M1, N0, N1, O0, O1, P0, and P1.
 1. Make sure the socket latches are open.
 2. Position the DIMM over the socket, making sure that the notch on the DIMM lines up with the key in the slot, then press the DIMM down into the socket until the side latches click in place.



3. Make sure that the latches are up and locked in place.
9. Install the three motherboard air baffles, replace the motherboard tray lid and then install the motherboard tray.

Refer to the instructions in the section *Reinstalling the Motherboard Tray*.
10. Connect all the cables to the motherboard tray.
11. Install all the power cords.
12. Power on the system and log in.
13. Confirm that the total memory is now 2 TB.

```
$ lsmem
```

```
Total online memory:      2T
```

14. Confirm that the system is healthy.

```
$ sudo nvsm show health
```

Chapter 14. Network Card Replacement

14.1. Network Card Replacement Overview

This is a high-level overview of the procedure to replace one or more network cards on the DGX A100 system.

1. Use the `nvsm show` commands to identify the failed network card.
2. Get a replacement card from NVIDIA Enterprise Support.
3. Shut down the system.
4. Label all motherboard tray cables and unplug them.
5. Remove the motherboard tray and open the lid.
6. Locate the failed network card and remove it.
7. Insert the new card into the slot and secure with the screw.
8. Close the lid on the motherboard tray, then insert the tray into the system.
9. Plug in all cables using the labels as a reference.
10. Power on the system.
11. Verify that the network card is healthy using `nvsm`.

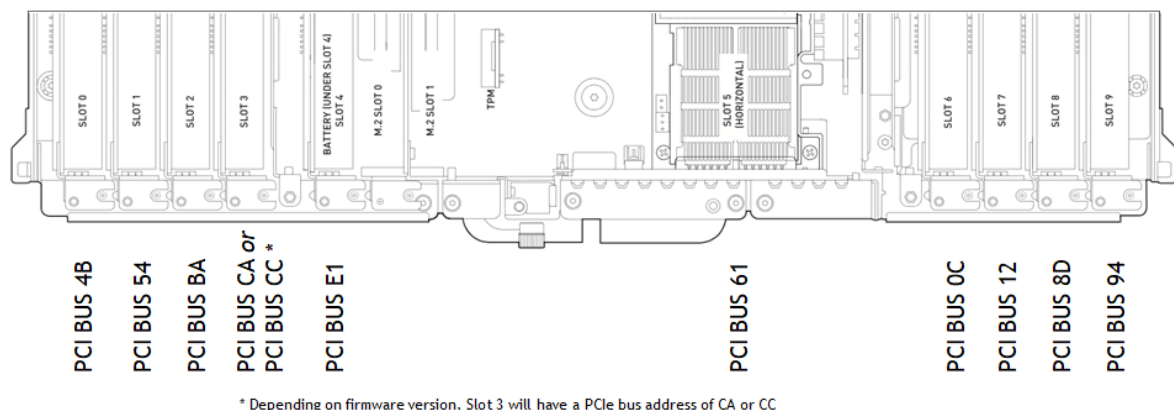
14.2. Identifying the Failed Network Card

Before attempting to replace any of the network cards, be sure to have performed the following:

1. Issue the following.

```
$ sudo nvsm show health
```

2. Match the PCIe bus ID for the failed card with the slot ID using the following diagram.



Note the slot ID for ordering and replacing.

- Order the appropriate card type as indicated in the following table, then follow the corresponding replacement instructions.

Table 1: Network Card Slot IDs

Slot ID	Card Type	Replacement Instructions
0-3, 6-9	Single port	<i>Replacing the Vertical Network Card</i>
4	Dual-port vertical	<i>Replacing the Vertical Network Card</i>
5	Dual-port horizontal	<i>Replacing the Horizontal Network Card</i>

14.3. Replacing the Vertical Network Card

Before attempting to replace any of the network cards, be sure to have performed the following:

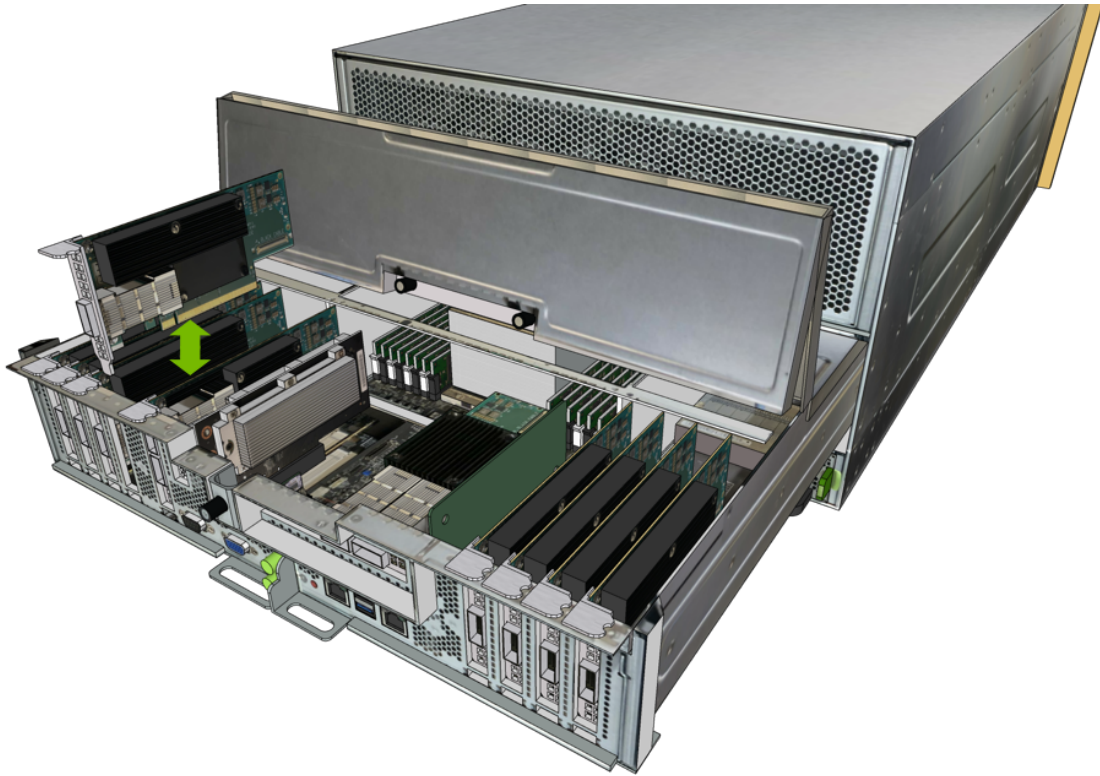
- Determined the location ID of the faulty network card needing replacement.
See *Identifying the Failed Network Card*.
- Obtained the replacement network card and have saved the packaging for use when returning the faulty component.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

- Power down the system.
- Label all network, monitor, and USB cables connected to the motherboard tray for easy identification when reconnecting.
- Unplug all power cords, and all network, monitor, and USB cables.
- Remove the motherboard tray.

Refer to the instructions in the section *Accessing the Motherboard Tray*.

5. Lift the network card off the motherboard and replace with the new network card.



6. Install the motherboard tray lid and then install the motherboard tray.
Refer to the instructions in the section [Replacing the Motherboard Tray](#).
7. Connect all the cables to the motherboard tray, connect all power cords, then power on the system and log in.
8. Confirm that the system is healthy.

```
$ sudo nvsm show health
```

There should be no new alerts listed.

Verify that the firmware is up to date according to the instructions in [Updating the Mellanox Network Card Firmware](#).

14.4. Replacing the Horizontal Network Card

Before attempting to replace any of the network cards, be sure to have performed the following:

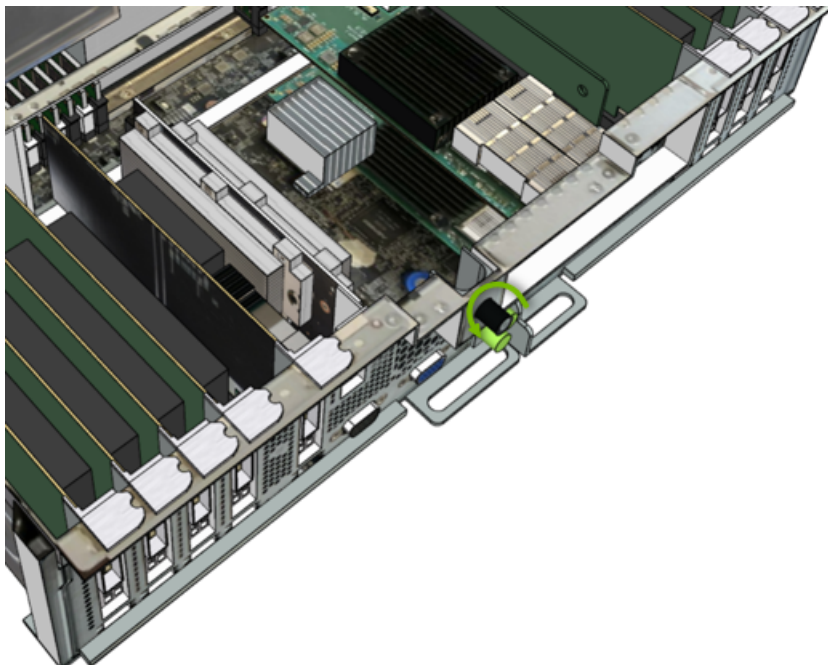
- Determined the location ID of the faulty network card needing replacement.
See [Identifying the Failed Network Card](#).
- Obtained the replacement network card have saved the packaging for use when returning the faulty component.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

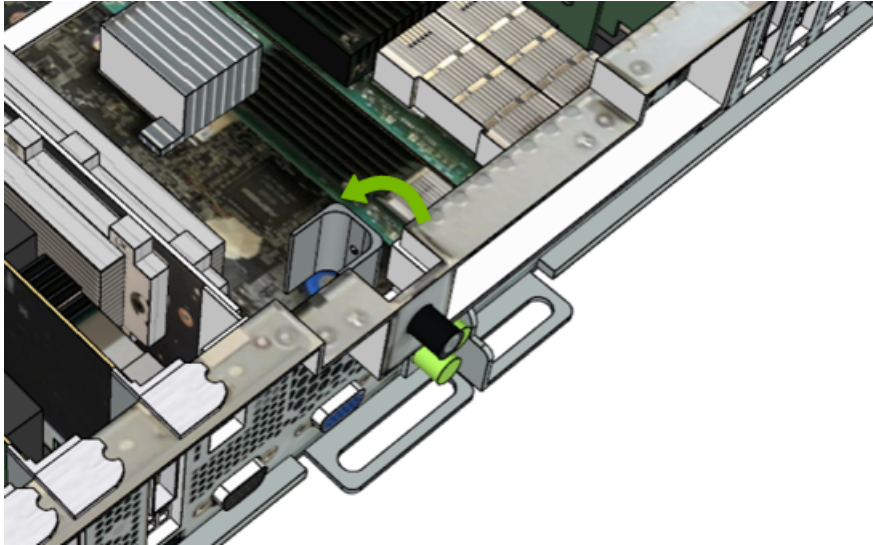
1. Power down the system.
2. Label all network, monitor, and USB cables connected to the motherboard tray for easy identification when reconnecting.
3. Unplug all power cords, and all network, monitor, and USB cables.
4. Remove the motherboard tray.

Refer to the instructions in the section [Accessing the Motherboard Tray](#).

5. Unlock the horizontal network card.
 1. Loosen the black thumbscrew that secures the PCIe card locking mechanism in place.

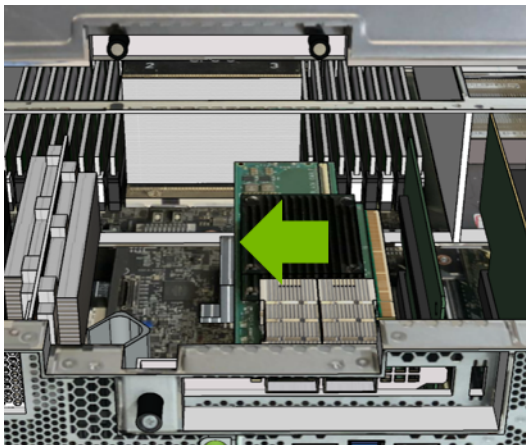


2. Open the locking mechanism by turning 90 degrees or more.

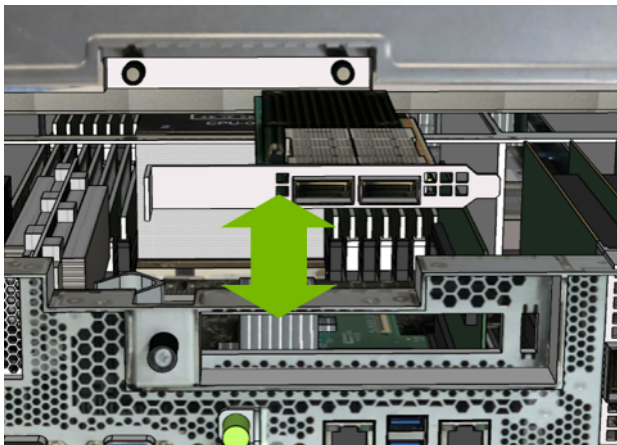


6. Replace the card.

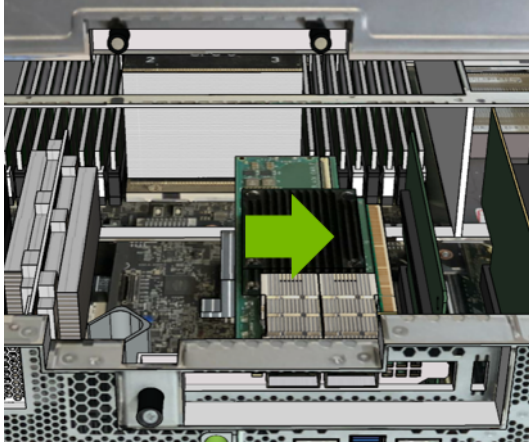
1. Pull the network card out of the riser card slot.



2. Replace the old network card with the new one.

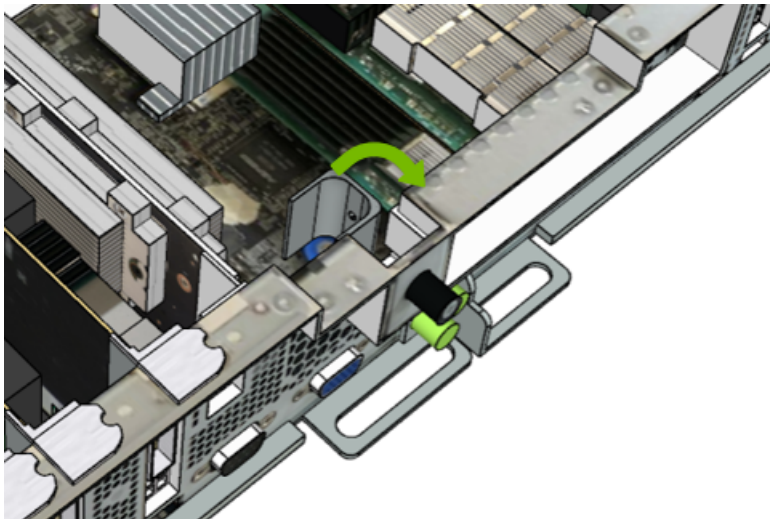


3. Install the network card into the riser card slot.

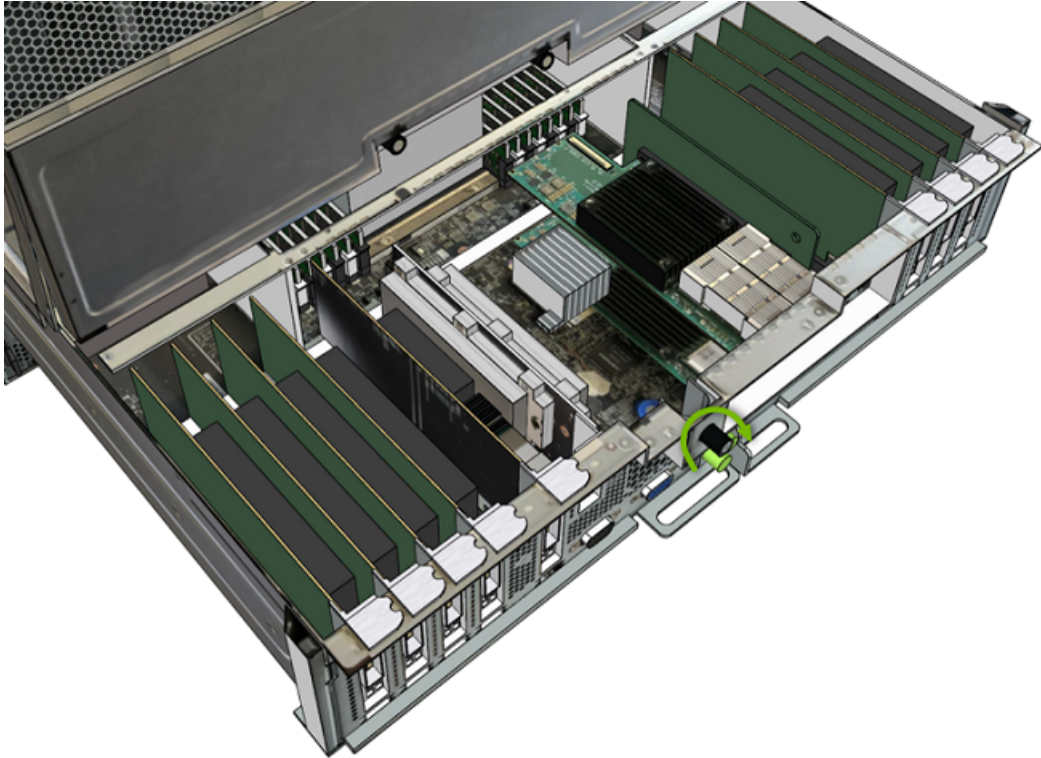


7. Lock the network card in place.

1. Close the locking mechanism by turning it back into its slot.



2. Tighten the black thumb screw to secure the card in place.



8. Install the motherboard tray lid and then install the motherboard tray.
Refer to the instructions in the section [Replacing the Motherboard Tray](#).
9. Connect all cables back into the network card ports.
10. Power on the system and log in.
11. Confirm that the system is healthy.

```
$ sudo nvsm show health
```

There should be no new alerts listed.

Verify that the firmware is up to date according to the instructions in [Updating the Mellanox Network Card Firmware](#).

Chapter 15. Adding the Optional Dual-port Horizontal Network Card

The DGX A100 comes with a vertical dual-port network card. You can expand the ports by adding a horizontal dual-port network card for slot 5.

15.1. Dual-port Network Card Upgrade Overview

This is a high-level overview of the procedure to install the optional horizontal dual-port network card on the DGX A100 system.

1. Obtain the new card from NVIDIA Sales.
2. Shut down the system.
3. Label all motherboard tray cables and unplug them.
4. Slide out the motherboard tray until it locks and then open the lid.
5. Remove the PCI slot filler for slot 5.
6. Insert the new card into the slot and secure with the screw.
7. Close the lid on the motherboard tray, then insert the tray into the system.
8. Plug in all cables using the labels as a reference.
9. Power on the system.
10. Verify that the network card is healthy using `nvsm show health`.

15.2. Adding the Horizontal Network Card

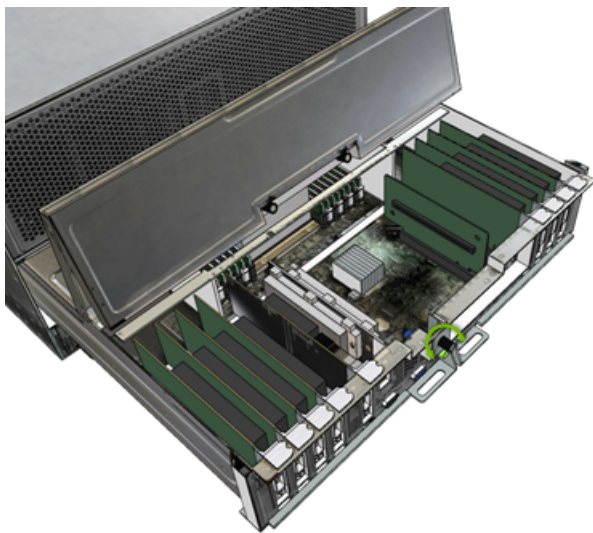
Be sure you have obtained the horizontal dual-port network card.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

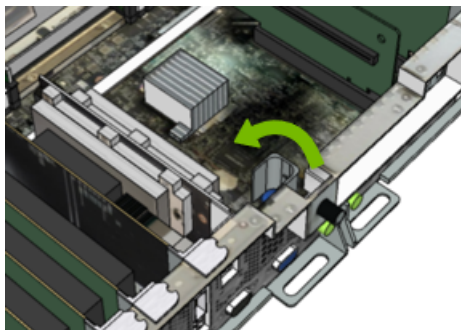
1. Power down the system.
2. Label all network, monitor, and USB cables connected to the motherboard tray for easy identification when reconnecting.
3. Unplug all power cords, and all network, monitor, and USB cables.
4. Remove the motherboard tray.

Refer to the instructions in the section [Accessing the Motherboard Tray](#).

5. Unlock the PCI EMI shield.
 1. Loosen the black thumbscrew that secures the locking mechanism in place.

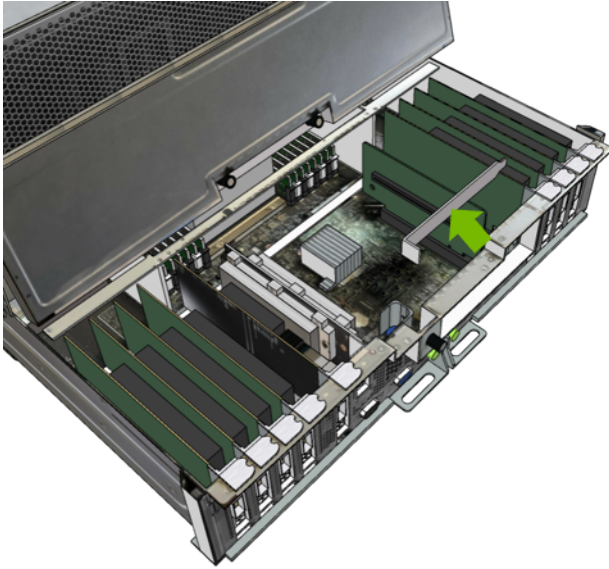


2. Open the locking mechanism by turning 90 degrees or more.

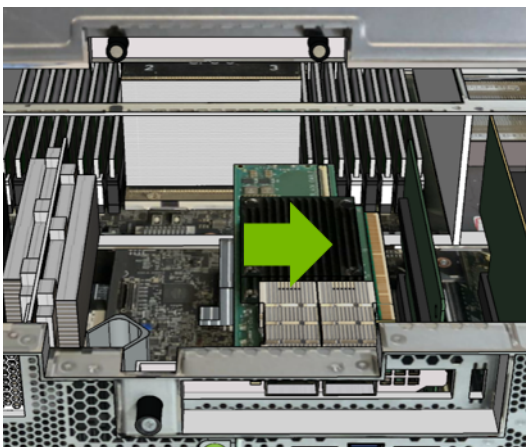
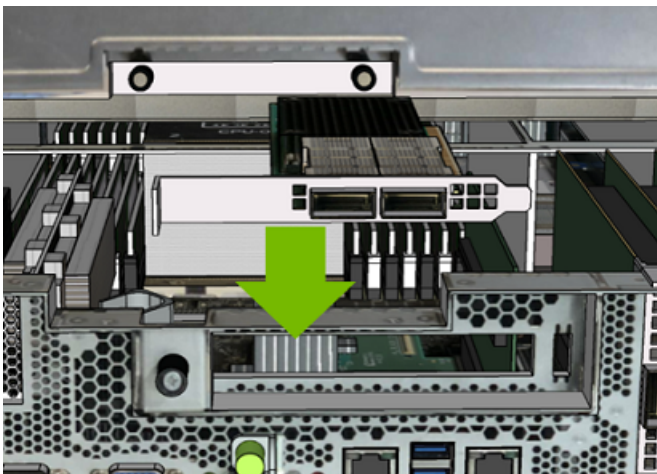


6. Remove the EMI shield.

Slide the EMI shield to the left to release it from the PCI slot on the riser card.

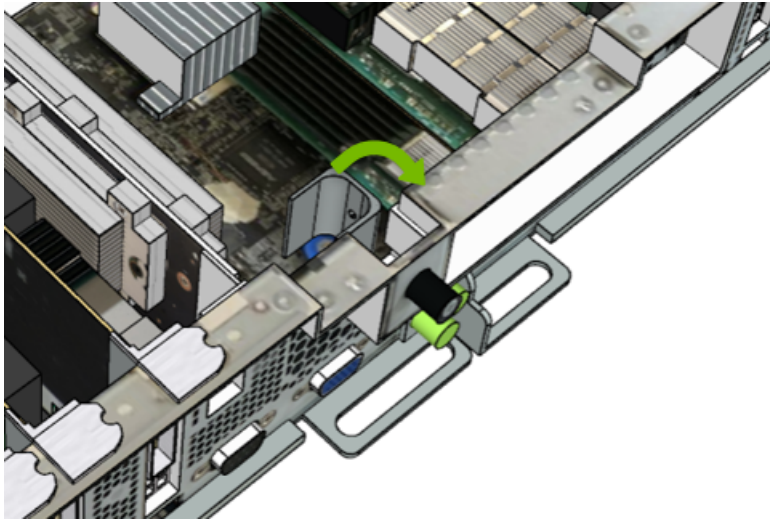


7. Install the new network card, inserting it into the slot on the riser card.

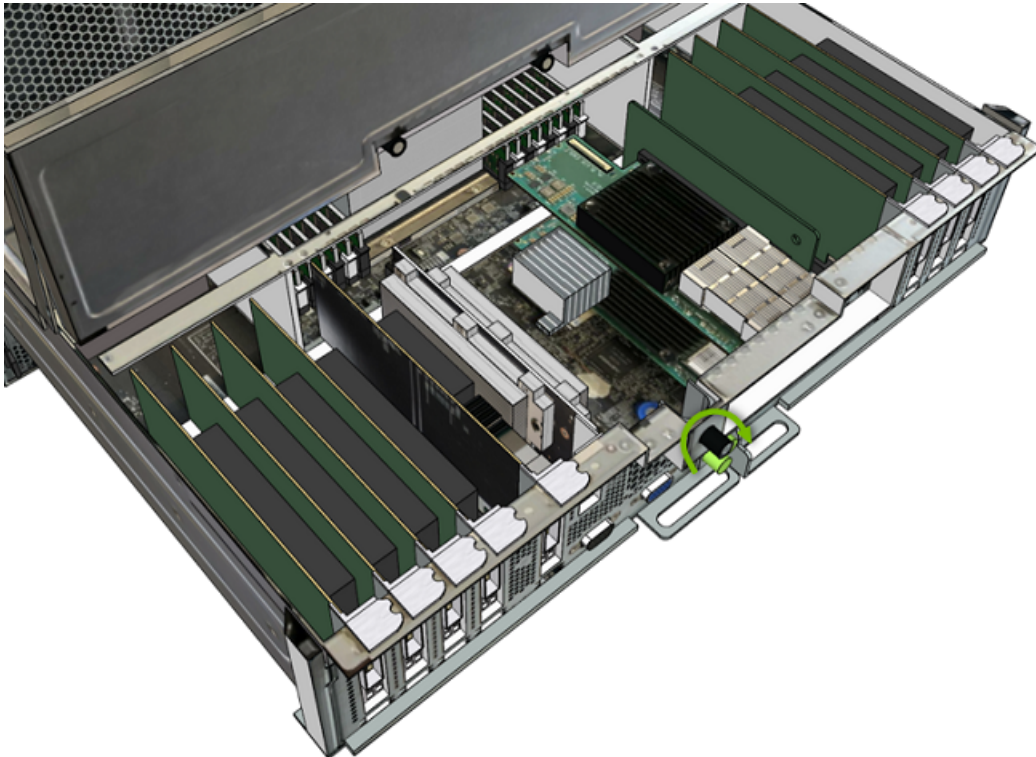


8. Lock the network card in place.

1. Close the locking mechanism by turning it back into its slot.



2. Tighten the black thumb screw to secure the card in place.



9. Install the motherboard tray lid and then install the motherboard tray.
Refer to the instructions in the section [Replacing the Motherboard Tray](#).
10. Connect all cables back into the network card ports.
11. Power on the system and log in.
12. Confirm that the system is healthy.

```
$ sudo nvsm show health
```

There should be no new alerts listed.

Verify that the firmware is up to date according to the instructions in [Updating the Mellanox Network Card Firmware](#).

Chapter 16. Updating the Mellanox Network Card Firmware

After replacing or installing the Mellanox ConnectX cards, make sure the firmware on the cards is up to date.

1. Confirm the OS is updated to the latest version; this will ensure the latest tested and supported firmware is downloaded.

```
$ sudo apt update
$ sudo apt full-upgrade
```

2. Run the firmware update script as follows:

```
$ sudo /opt/mellanox/mlnx-fw-updater/mlnx_fw_updater.pl
```

If the firmware version in the new card is more recent than the one available in the operating system, the firmware should be downgraded to a supported version by issuing the following:

```
$ sudo /opt/mellanox/mlnx-fw-updater/mlnx_fw_updater.pl --force-fw-update
```

3. Reboot the system for the firmware update to take effect.

```
$ sudo reboot
```

4. Once the system comes up, make sure the firmware versions are all the same by issuing:

```
$ cat /sys/class/infiniband/mlx5_*/fw_ver
```

Chapter 17. Front Console Board Replacement

17.1. Front Console Board Replacement Overview

This is a high-level overview of the procedure to replace the front console board on the DGX A100 system.

1. Unpack the new front console board.
2. Shut down the system.
3. Using a Phillips #2 screwdriver, release the captive screws.
4. Pull the front console board out of the system.
5. Insert the new front console board.
6. Tighten the screws.
7. Power on the system and confirm the ports work.
8. Ship the failed unit back to NVIDIA Enterprise Support using the provided packaging.

17.2. Replacing the Front Console Board

A front console board malfunction can be determined in a few ways.

- ▶ No display or connectivity occurring after plugging in a keyboard and monitor to the front of the system, or
- ▶ The USB ports do not work, or
- ▶ The front temperature sensor does not provide a temperature reading.

Raise a ticket with NVIDIA Enterprise Services to request a replacement.

When the new board arrives, unpack it and keep the packaging to use for sending back the old board.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

1. Power down the system.
2. Remove the bezel.
3. Replace the front console board.
 1. Using a Phillips #2 screwdriver, loosen the two captive screws that secure the front console board.



2. Replace the front console board.



3. Tighten the screws.



4. Confirm functionality.

1. Power on the system.
2. Issue the following to confirm the temperature sensor is working properly.

```
$ sudo nvsm show health
```

5. Return the old module to NVIDIA Enterprise Services.

Chapter 18. Motherboard Tray Battery Replacement

18.1. Motherboard Tray Battery Replacement Overview

This is a high-level overview of the procedure to replace the DGX A100 system motherboard tray battery.

1. Get a replacement battery - type CR2032.
2. Shut down the system.
3. Label all motherboard cables and unplug them.
4. Slide out the motherboard tray.
5. Open the motherboard tray IO compartment.
6. Pull out the M.2 riser card with both M.2 disks attached.
7. Pull out the dual port Mellanox card (slot 4).
8. Replace the battery on the motherboard.
9. Install the dual port Mellanox card (slot 4).
10. Install the M.2 riser card with both M.2 disks attached.
11. Close the lid IO compartment on the motherboard tray.
12. Slide the motherboard tray into the system.
13. Plug in all cables using the labels as a reference.
14. Power on the system.
15. Confirm the system is healthy by running `nvsm show health`

18.2. Replacing the Motherboard Tray Battery

A battery failure can be determined in a few ways.

- ▶ “Invalid configuration” will appear on your screen.
- ▶ Setup appears on your screen before booting.
- ▶ “Press F1 to continue” appears on the console.
- ▶ A Clock Error or Clock Message appears on your screen.
- ▶ The system clock loses time and date.

Call NVIDIA Enterprise Support to confirm that the battery is the right component to replace. The CR2032 battery is not provided by NVIDIA, but can be purchased from a convenience store.

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

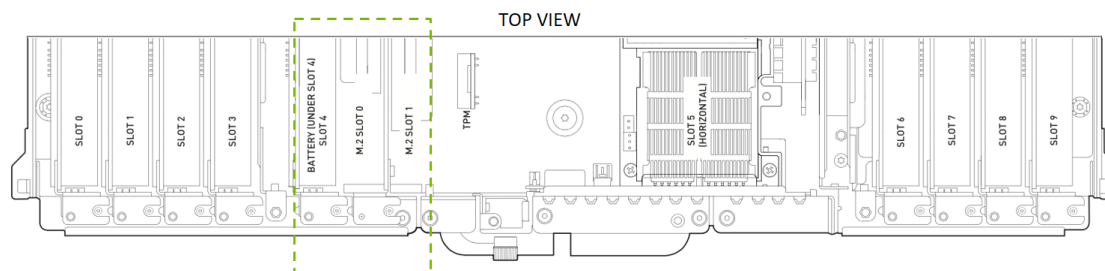
1. Power down the system.
2. Label all cables connected to the motherboard tray for easy identification when reconnecting.
3. Unplug the cables.
4. Remove the motherboard tray.

Refer to the instructions in the section [Accessing the Motherboard Tray](#).

5. Remove the M.2 riser card, the IO card, and the air baffle.

These components need to be removed to gain access to the battery.

1. Confirm the locations of the components to remove using the following diagram.

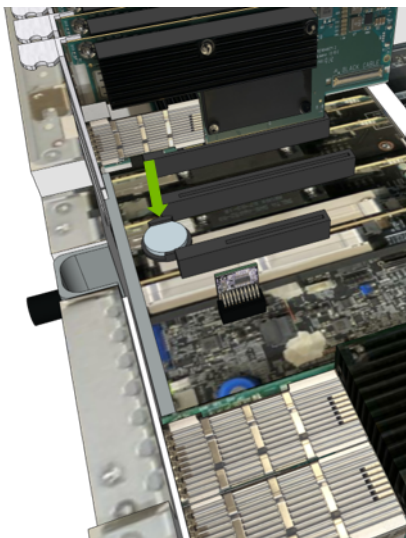


2. Lift off the IO card, the M.2 riser, and the air baffle and set them on a solid surface.

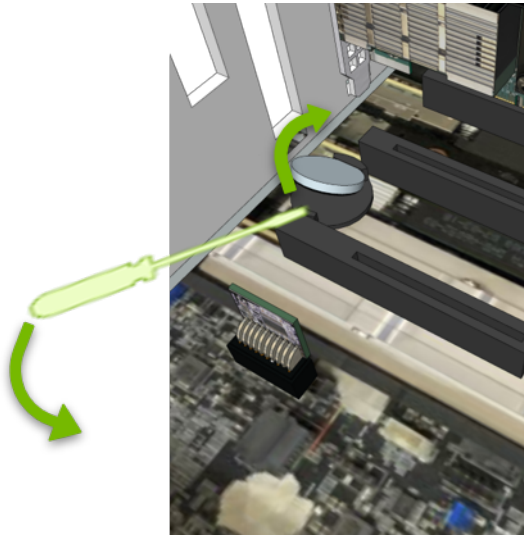


6. Replace the battery.

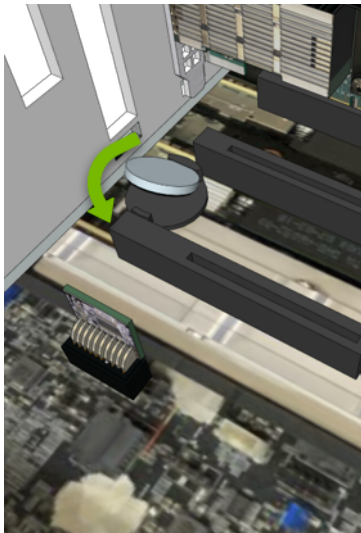
1. Locate the battery, using the following image as a guide.



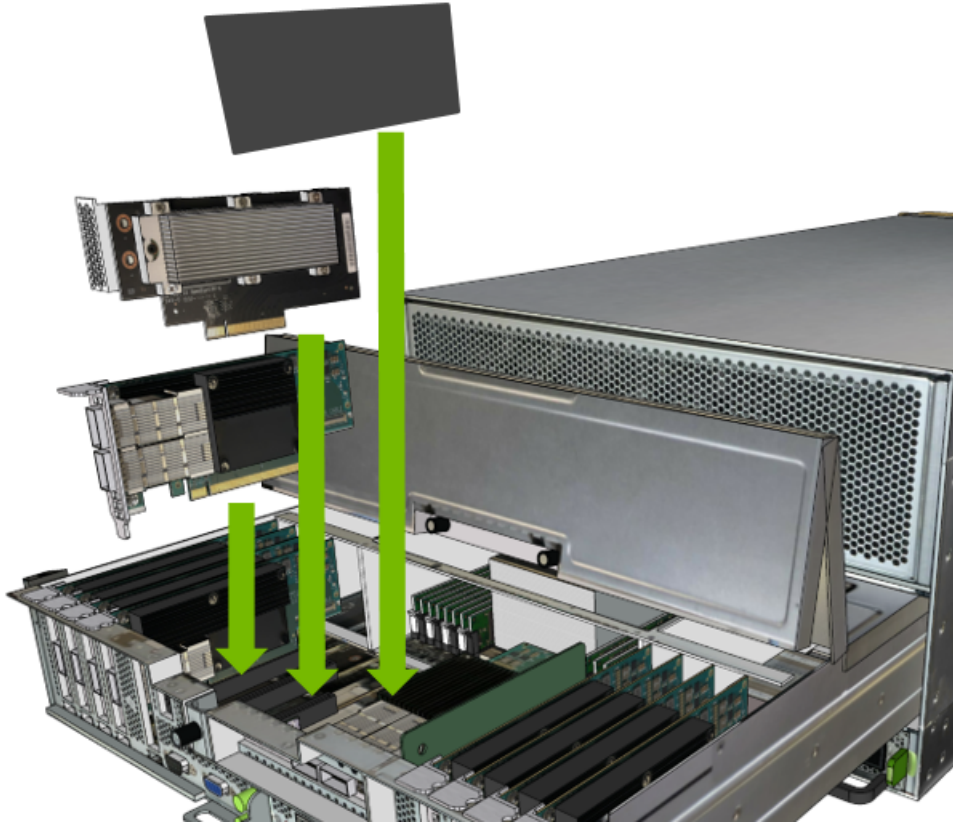
2. Use a small flat-head screwdriver or similar thin tool to gently lift the battery from the battery holder.



3. Replace the battery with a new CR2032, installing it in the battery holder.



7. Re-insert the IO card, the M.2 riser card, and the air baffle into their respective slots.



8. Replace the motherboard tray.

Refer to the instructions in the section [Replacing the Motherboard Tray](#).

9. Connect all the cables and power cords to the motherboard tray.
10. Apply power to the system and then log in.
11. Restore the date on the system.

1. Set the date.

```
$ sudo date [MMDDhhmm[[CC]YY][.ss]]
```

2. Sync the date and time to the hardware real time clock.

```
$ sudo hwclock -w
```

3. Reset the BMC

```
$ sudo ipmitool mc reset cold
```

12. Confirm that the time and date on the system are updated.

```
$ sudo nvsm show health
```

Chapter 19. Trusted Platform Module Replacement

19.1. Trusted Platform Module Replacement Overview

This is a high-level overview of the procedure to replace the trusted platform module (TPM) on the DGX A100 system.

1. If enabled, disable drive encryption.
2. Shut down the system.
3. Label all motherboard tray cables and unplug them.
4. Slide out the motherboard tray and open the motherboard tray I/O compartment.
5. Replace the TPM on the motherboard.
6. Close the I/O compartment lid on the motherboard tray.
7. Slide the motherboard tray into the system.
8. Plug in all cables using the labels as a reference.
9. Power on the system.
10. Verify that the network card is healthy using `nvsm show health`.
11. If the data drives need to be protected, then enable encryption.

19.2. Replacing the Trusted Platform Module

Caution: Static Sensitive Devices: - Be sure to observe best practices for electrostatic discharge (ESD) protection. This includes making sure personnel and equipment are connected to a common ground, such as by wearing a wrist strap connected to the chassis ground, and placing components on static-free work surfaces.

1. Obtain a new Trusted Platform Module (TPM) from NVIDIA.

2. If data drives are encrypted, then disable encryption.

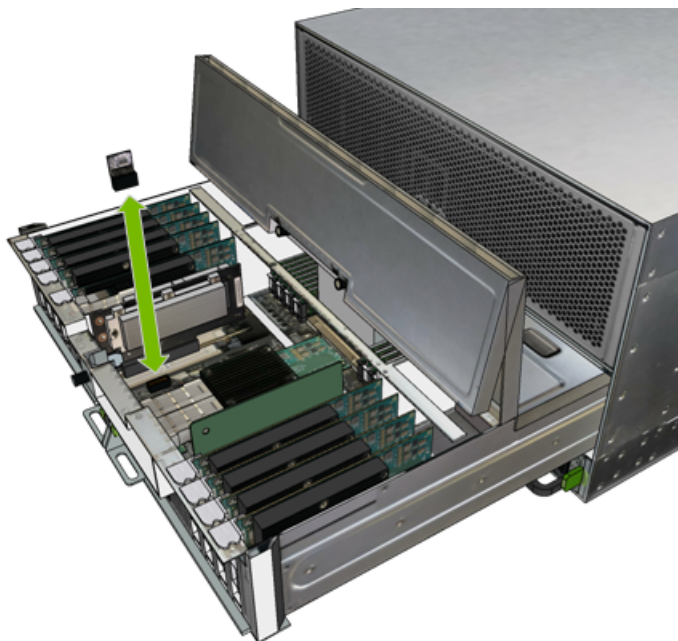
```
$ sudo nv-disk-encrypt disable
```

Note: The TPM2 OS package must be installed and TPM enabled in the SBIOS. Refer to the chapter *Managing the DGX A100 Self-Encrypting Drives* in the [NVIDIA DGX A100 User Guide](#) for more information.

3. Power down the system.
4. Label all cables connected to the motherboard tray for easy identification when reconnecting.
5. Unplug all power cords.
6. Unplug all network, monitor, and USB cables.
7. Remove the motherboard tray.

Refer to the instructions in the section [Removing the Motherboard Tray](#).

8. Replace the TPM by pulling it up vertically and then inserting the replacement in its slot.



9. Close the motherboard tray lid and reinsert the mother board tray and secure.
Refer to the instructions in the section [Reinstalling the Motherboard Tray](#).
10. Connect all the cables to the motherboard tray.
11. Install all the power cords.
12. Power on the system and log in.
13. If data drives were encrypted prior to replacing the TPM, encrypt the drives.

Note: The TPM2 OS package must be installed and TPM enabled in the SBIOS. Refer to the chapter *Managing the DGX A100 Self-Encrypting Drives* in the [NVIDIA DGX A100 User Guide](#) for

more information.

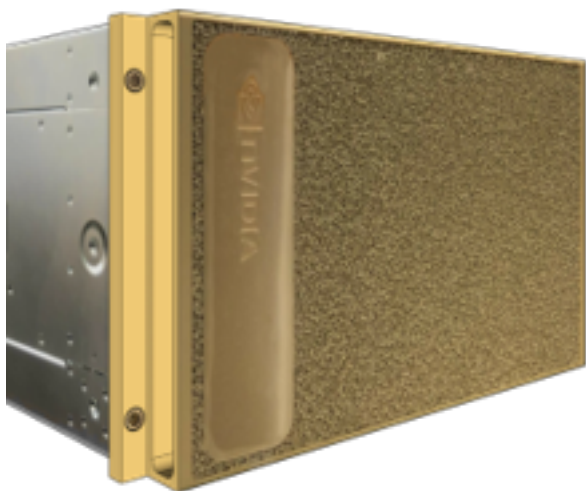
The following is an example command for enabling drive encryption:

```
$ sudo nv-disk-encrypt init -g -r -k <your vault password>
```

14. Confirm that the system is healthy.

```
$ sudo nvsm show health
```

Chapter 20. Removing and Attaching the Bezel

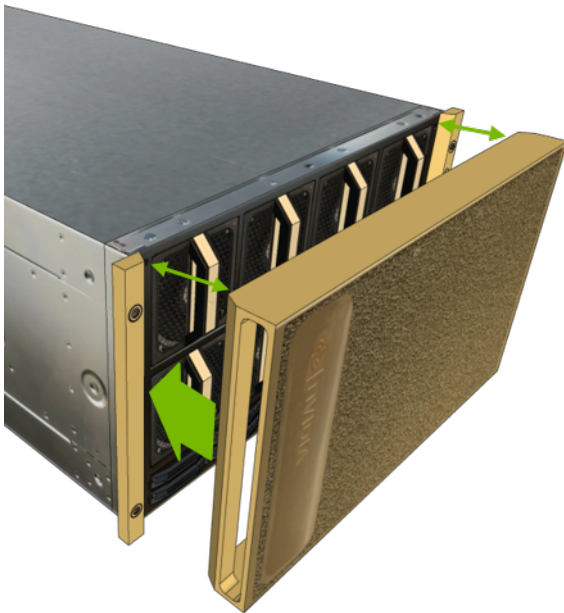


1. Grab the bezel on both sides by the side handles, then pull directly away from the system to disengage from the magnetic latch.



2. To replace the bezel, align the bezel alignment pins with the chassis, then let the magnetic latch

complete the attachment of the bezel.





Chapter 21. Installing the Rack Mount Kit

21.1. Installing the Rails

Follow these instructions to install the DGX A100 server rack mount kit. The rack mount kit acts as a shelf in the rack, it does not allow the system to be moved once installed. All components are serviceable from the front or rear, so this movement is not necessary.

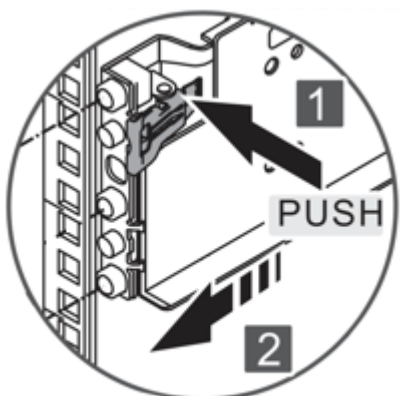


Item	Description
1	Bottom lip to support the DGX A100.
2	Prongs that enter the holes of the rack. Holes can be square or round.
3	Metal clips that hold the rail in place.
4	<p>Thread for securing the rack mount kit to the rack. Two types of screws are provided.</p> <ul style="list-style-type: none">► Flat head: Must be used along with the provided washers at the front of the rack so that the DGX A100 can be installed flush to the rack.  <ul style="list-style-type: none">► Pan or round head: Must be used at the rear of the rack. 

- Use a Phillips screwdriver to assist in mounting the rails to the rack.

- Follow any designations on the slide rail to determine front/back and left-side/right-side positioning against the rack.
- 1. Align the bottom lip of the left or right rail to the bottom of the first rack unit for the server.
- 2. Attach the front of the rail to the rack.

Push the metal tab on the rail and then insert the two spring-loaded prongs into the holes on the front rack post.



FRONT LEFT

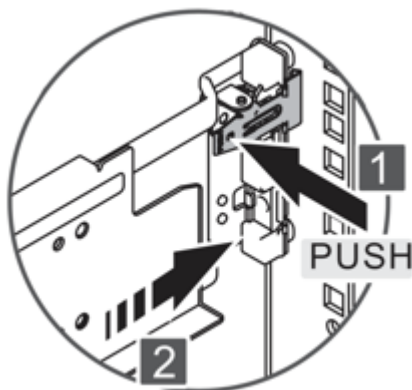
On square-holed racks, make sure the prongs are completely inserted into the hole by confirming that the spring is fully extended.

The metal clip helps secure the rail in place while installing the rail to the rear of the rack.

- 3. Extend the rail to the rear post of the rack and secure the rail to the post.

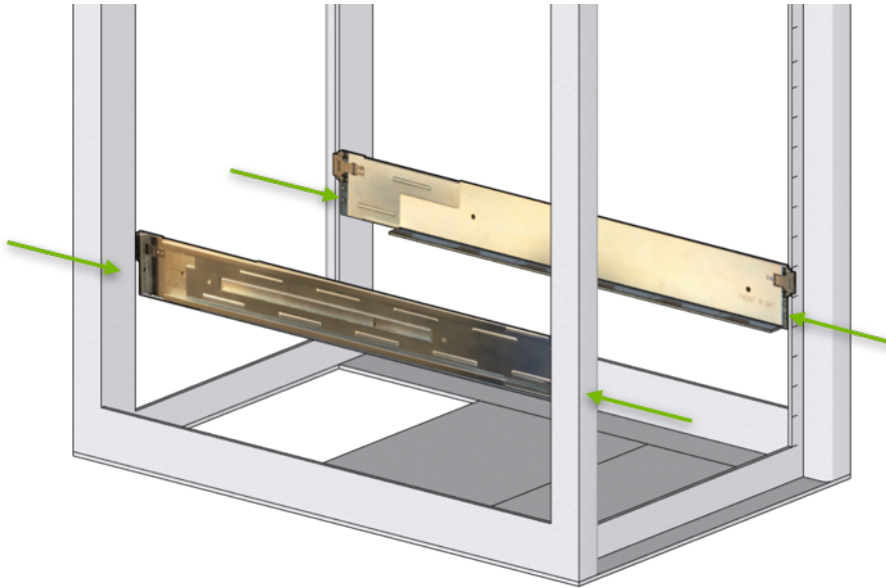
Make sure that rail is level and the attachment on the rear post is at the same rack unit as the front.

- 1. Insert the spring-loaded prongs into the holes on the rear rack post.



- 2. On square-holed racks, make sure the prongs are completely inserted into the hole by confirming that the spring is fully extended.
- 4. Repeat these steps for the other rail.
- 5. Secure the rails to the rack using the provided screws.

Install the four screws in the bottom holes of the rack kit.



Do not use the top holes, as those will be used for the captive screw in the chassis.

6. Confirm that the rails have been installed properly.
 - ▶ Both sides are installed on the same rack unit and are horizontal to each other.
 - ▶ The bottom lip is at the same height on all four posts.
 - ▶ The metal clips are properly attached.
 - ▶ Four screws are installed - flat head on the front and pan head on the back.

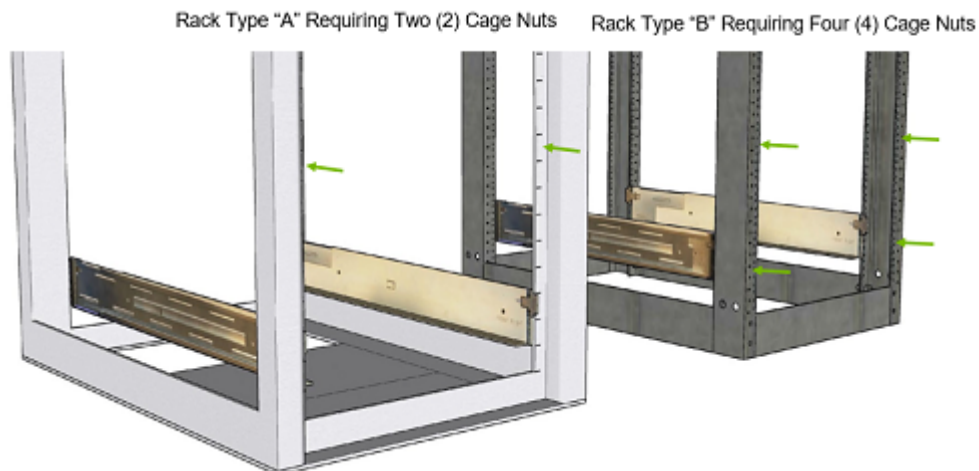
21.2. Installing the Cage Nuts

The DGX A100 server is secured to the rack using four captive screws - one at each corner of the front of the unit.

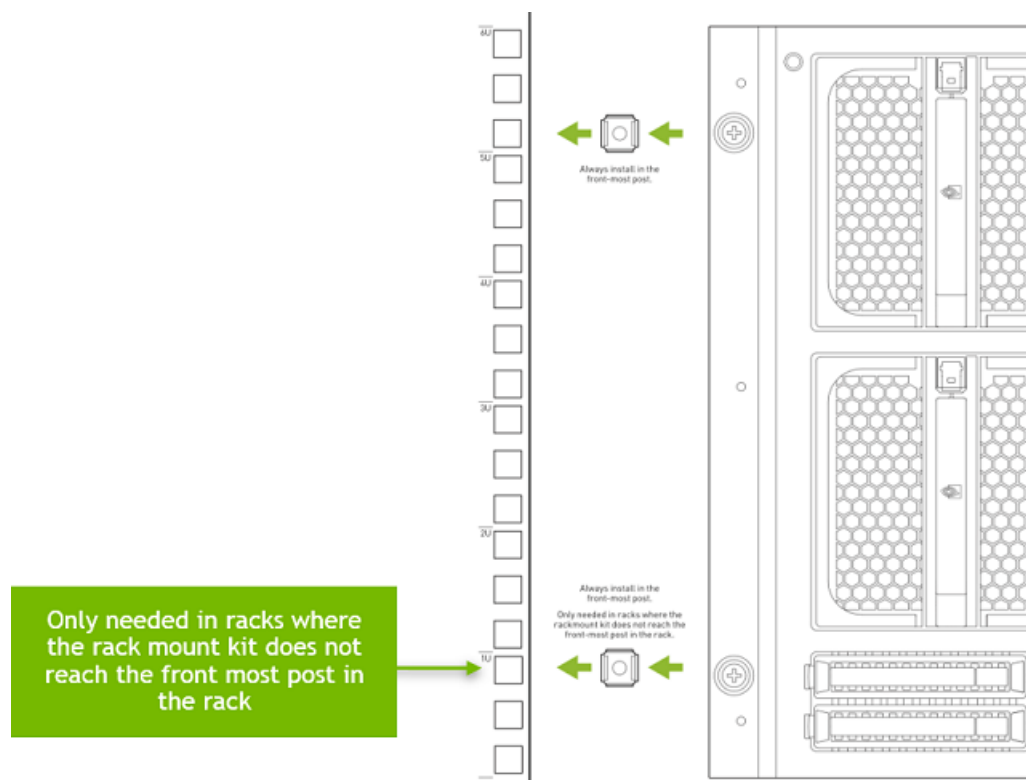
- ▶ If your rack has round holes with 10-32 threads, then the screws will attach directly to the rack mounting holes.
- ▶ If your rack has square holes, then you need to install the included cage nuts to provide attachment points for mounting the DGX A100 server to the front of the rack.

The number of cage nuts to install depends on the type of rack that is used. On most racks, cage nuts are needed only for the top screws. The bottom screws will attach to the threads in the rails. The image on the left side of the following diagram shows this type. Green arrows indicate the general area for the cage nuts.

The image on the right side shows a rack with square posts such that the rail does not meet the front-most part of the rack. In this rack, cage nuts need to be installed at the bottom holes as well, as indicated by the green arrows. All four cage nuts are also needed if you have installed a rack shelf instead of rack rails.



Use the provided template to determine the exact location for installing the cage nuts.



1. Place the template so that the bottom of the template rests on the rail lip (or at the same level as the rail lip).

The template is double sided so it can be used as a reference on the left and right posts of the rack.

2. Install the cage nuts at the positions indicated on the template.
 - Rail kits attached to Type A racks require two (2) cage nuts installed; top positions only.
 - Rail kits attached to Type B racks require four (4) cage nut installed; both top and bottom positions.

Chapter 22. Third-Party License Notices

This NVIDIA product contains third party software that is being made available to you under their respective open source software licenses. Some of those licenses also require specific legal information to be included in the product. This section provides such information.

22.1. Micron msecli

The `msecli` utility is provided under the following terms:

Micron Technology, Inc. Software License Agreement PLEASE READ THIS LICENSE AGREEMENT ("AGREEMENT") FROM MICRON TECHNOLOGY, INC. ("MTI") CAREFULLY: BY INSTALLING, COPYING OR OTHERWISE USING THIS SOFTWARE AND ANY RELATED PRINTED MATERIALS ("SOFTWARE"), YOU ARE ACCEPTING AND AGREEING TO THE TERMS OF THIS AGREEMENT. IF YOU DO NOT AGREE WITH THE TERMS OF THIS AGREEMENT, DO NOT INSTALL THE SOFTWARE. LICENSE: MTI hereby grants to you the following rights: You may use and make one (1) backup copy the Software subject to the terms of this Agreement. You must maintain all copyright notices on all copies of the Software. You agree not to modify, adapt, decompile, reverse engineer, disassemble, or otherwise translate the Software. MTI may make changes to the Software at any time without notice to you. In addition MTI is under no obligation whatsoever to update, maintain, or provide new versions or other support for the Software. OWNERSHIP OF MATERIALS: You acknowledge and agree that the Software is proprietary property of MTI (and/or its licensors) and is protected by United States copyright law and international treaty provisions. Except as expressly provided herein, MTI does not grant any express or implied right to you under any patents, copyrights, trademarks, or trade secret information. You further acknowledge and agree that all right, title, and interest in and to the Software, including associated proprietary rights, are and shall remain with MTI (and/or its licensors). This Agreement does not convey to you an interest in or to the Software, but only a limited right to use and copy the Software in accordance with the terms of this Agreement. The Software is licensed to you and not sold.

DISCLAIMER OF WARRANTY:

THE SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. MTI EXPRESSLY DISCLAIMS ALL WARRANTIES EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, NON-INFRINGEMENT OF THIRD PARTY RIGHTS, AND ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. MTI DOES NOT WARRANT THAT THE SOFTWARE WILL MEET YOUR REQUIREMENTS, OR THAT THE OPERATION OF THE SOFTWARE WILL BE UNINTERRUPTED OR ERROR-FREE. FURTHERMORE, MTI DOES NOT MAKE ANY REPRESENTATIONS REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE IN TERMS OF ITS CORRECTNESS, ACCURACY, RELIABILITY, OR OTHERWISE. THE ENTIRE RISK ARISING OUT OF USE OR PERFORMANCE OF THE SOFTWARE REMAINS WITH YOU. IN NO EVENT SHALL MTI, ITS AFFILIATED COMPANIES OR THEIR SUPPLIERS BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL, INCIDENTAL, OR SPECIAL DAMAGES (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF PROFITS, BUSINESS INTERRUPTION, OR LOSS OF

INFORMATION) ARISING OUT OF YOUR USE OF OR INABILITY TO USE THE SOFTWARE, EVEN IF MTI HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Because some jurisdictions prohibit the exclusion or limitation of liability for consequential or incidental damages, the above limitation may not apply to you.

TERMINATION OF THIS LICENSE: MTI may terminate this license at any time if you are in breach of any of the terms of this Agreement. Upon termination, you will immediately destroy all copies the Software.

GENERAL: This Agreement constitutes the entire agreement between MTI and you regarding the subject matter hereof and supersedes all previous oral or written communications between the parties. This Agreement shall be governed by the laws of the State of Idaho without regard to its conflict of laws rules.

CONTACT: If you have any questions about the terms of this Agreement, please contact MTI's legal department at (208) 368-4500. By proceeding with the installation of the Software, you agree to the terms of this Agreement. You must agree to the terms in order to install and use the Software.

22.2. Mellanox (OFED)

MLNX_OFED <<http://www.mellanox.com/>> is provided under the following terms:

Copyright (c) 2006 Mellanox Technologies. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Chapter 23. Notices

23.1. Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation (“NVIDIA”) makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer (“Terms of Sale”). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer’s own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or

services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

23.2. Trademarks

NVIDIA, the NVIDIA logo, DGX, DGX-1, DGX-2, DGX A100, DGX Station, and DGX Station A100 are trademarks and/or registered trademarks of NVIDIA Corporation in the United States and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

©2022-2024, NVIDIA Corporation