Configuring the NVIDIA ConnectX-8 Adapter and Auxiliary Cable on ThinkSystem Servers

Planning / Implementation

The NVIDIA ConnectX-8 800Gb adapters require a PCIe Gen6 x16 or PCIe Gen5 x32 host interface to maximize performance of the adapter. On PCIe Gen5 servers such as the ThinkSystem V4 servers, a x32 host interface is achieved through the use of both the x16 interface of the adapter's edge connector plus a x16 auxiliary cable which connects to an adjacent slot in the server.

By default, the ConnectX-8 adapters are configured to only use the x16 host interface of the adapter itself. In order to use the auxiliary cable, settings in the adapter will need to be changed. This document explains the steps how to configure the adapter operating mode, so that the x16 connector of the auxiliary cable is also enabled.

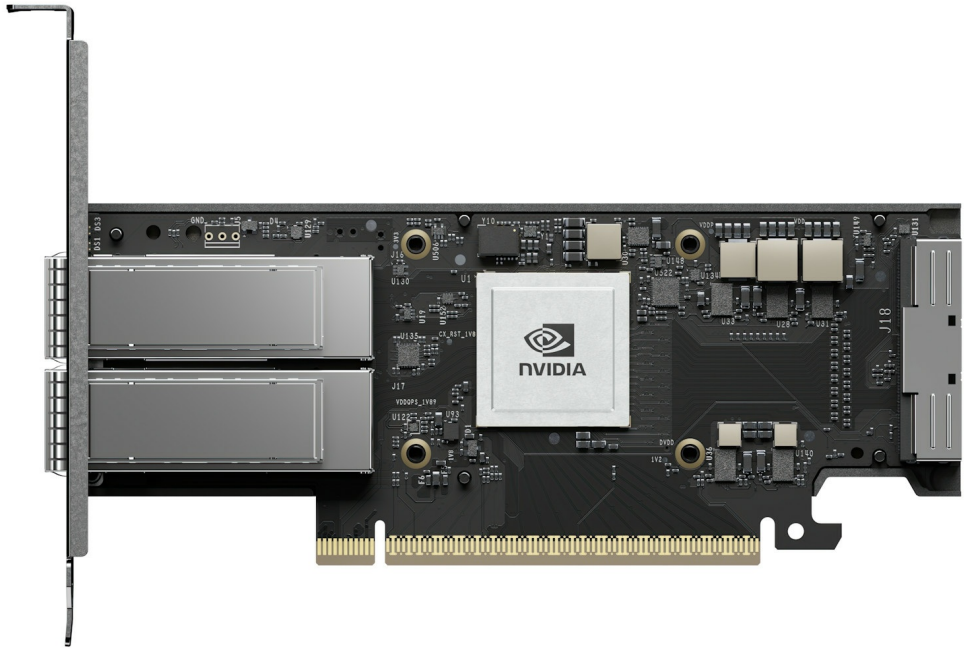

Figure 1. ThinkSystem NVIDIA ConnectX-8 8240 Adapter with x16 edge connector and x16 auxiliary cable connector

The default Gen6 x16 mode of the adapter does not auto-negotiate to Gen5 x32 due to firmware constraints because the Socket Direct function is not able to automatically detect when the Auxiliary cable is connected. As a result, manual configuration is required, as documented in this paper.

Socket Direct: If you are planning to use the Socket Direct of the adapters, you will also need to follow these instructions to enable the Gen5 x32 host connectivity. Socket Direct enables direct access from two CPUs, each with a x16 host interface. One processor connects through the x16 edge connector of the adapter and the other processor connects through a x16 auxiliary cable connected to the adapter.

Applicable adapters

This document applies to the adapters and auxiliary cable listed in the following table

Table 1. Applicable adapters and auxiliary cable

Part number	Feature code	Description
ConnectX-8 adapters		
4XC7B03668	C9AQ	ThinkSystem NVIDIA ConnectX-8 8240 400GbE / 400Gb/s IB QSFP112 2-port PCIe Gen6 x16 (Generic FW)
4XC7B03667	C9AP	ThinkSystem NVIDIA ConnectX-8 8180 800Gbs XDR IB / 2x400GbE OSFP 1-port PCIe Gen6 x16 (Generic FW)
CTO only	CAGC	ThinkSystem NVIDIA ConnectX-8 8180 800Gbs XDR IB / 2x400GbE OSFP 1-Port PCIe Gen6 x16 DWC (Generic FW)
Auxiliary cable		
4X97B05994	C9NX	ThinkSystem 1U/2U V4 NVIDIA ConnectX-8 Aux Cable Kit

Prerequisites

Ensure that the adapter and cable are installed in supported slots. Consult the server product guide to verify:

A list of product guides of supported servers can be found in the following link:
https://lenovopress.lenovo.com/search#term=8180%2520or%25208240&resource_type=product-guide&sort=relevance

Installing the tools

Configuring the CX8 operating mode requires the use of NVIDIA's command-line tools, `mst` and `mlxconfig`. Both commands are required for this procedure, and are part of the NVIDIA DOCA downloadable package.

To install NVIDIA DOCA, perform the following steps.

1. Download the DOCA package from: https://developer.nvidia.com/doca-downloads?deployment_platform=Host-Server&deployment_package=DOCA-Host
2. Make the selections on the web page appropriate for your server:

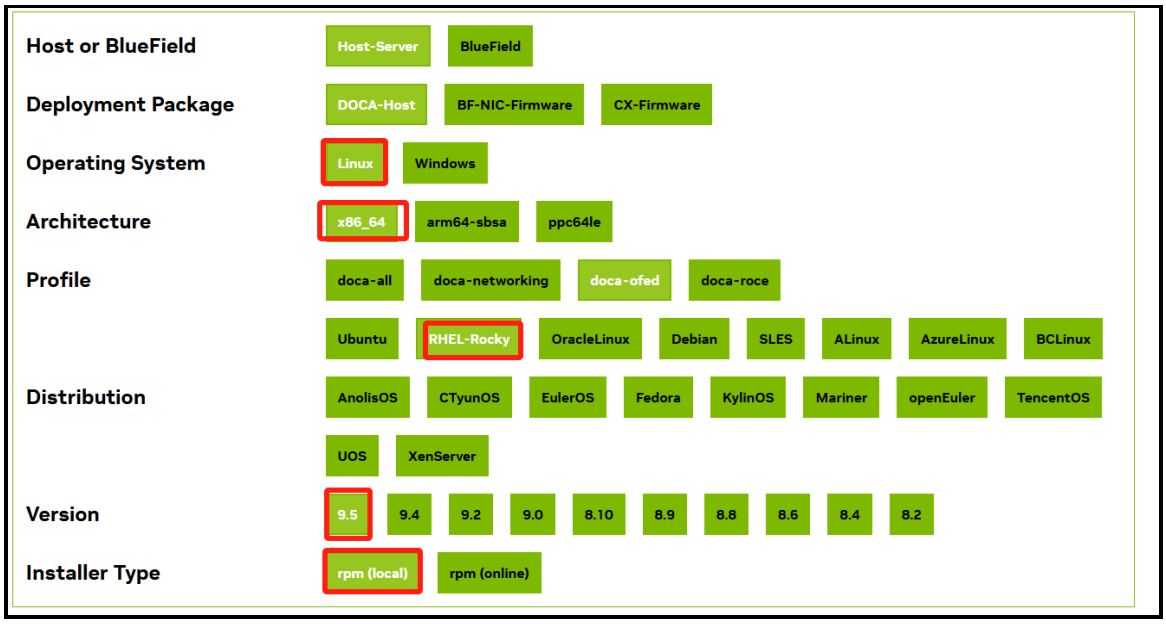

Figure 2. DOCA download page selections

3. You'll then be prompted for installation instructions, similar to the following figure.

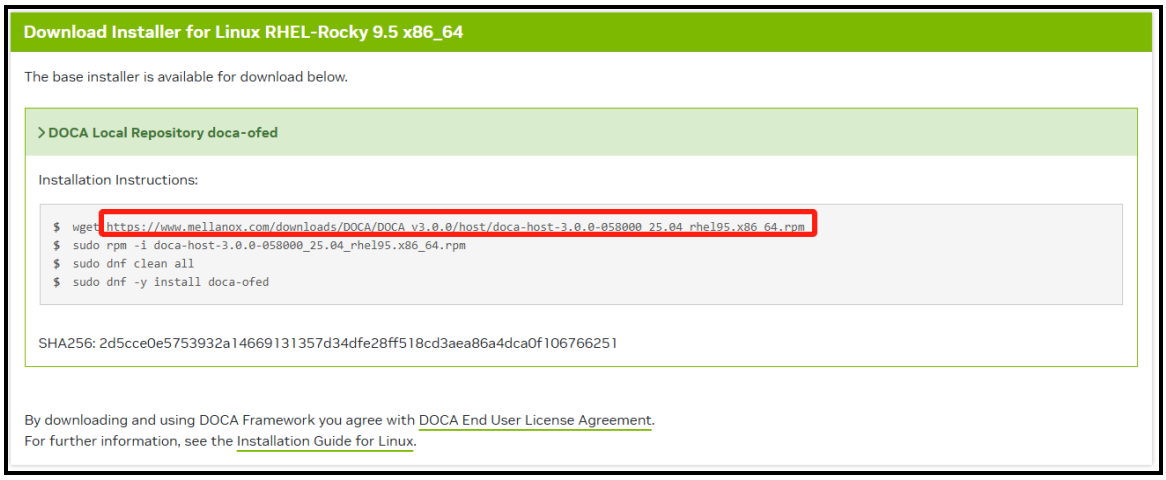

Figure 3. DOCA installation instructions

Parameters of the mlxconfig command

In the next section, we'll be using the `mlxconfig` command to configure the adapter to operate in PCIe Gen5 x32 mode. The commands we'll use are the following:

```
# mlxconfig -d devicename -y set PCIE_CREDIT_TOKEN_TIMEOUT=0
# mlxconfig -d devicename -y set PCI_BUS00_WIDTH=5
# mlxconfig -d devicename -y set PCI_BUS00_SPEED=4
# mlxconfig -d devicename -y set PCI_BUS00_HIERARCHY_TYPE=0
# mlxconfig -d devicename -y set PCI_BUS10_WIDTH=5
# mlxconfig -d devicename -y set PCI_BUS10_SPEED=4
# mlxconfig -d devicename -y set PCI_BUS10_HIERARCHY_TYPE=0
```

In these commands, `device name` is the name of configuration node for the CX8 adapter. In our example, as shown in the `mst status` command, is **`/dev/mst/mt4131_pciconf0`**.

The values after the **`-y set`** parameter are used to configure the adapter, and are listed in the following table.

Table 2. Values of the `-y set` parameter

Parameter	Description	Values
PCIE_CREDIT_TOKEN_TIMEOUT	PCIe credit timeout	0x0 = the device will adjust the timeout 0xffff = disables the feature
PCI_BUS<XY>	Adapter's PCIe lanes The CX8 chip has 48 PCIe lanes, divided into 3 interfaces (X=0/1/2); Each interface has 16 lanes, divided into 8 groups (Y= 0-7);	X = 0/1/2 0 = CX8 adapter 1 = Aux cable 2 = unused Y = 0-7 PCI_BUS0Y refers to the lanes of the CX8 adapter; there are 8 pairs in total. PCI_BUS1Y refers to the lanes of the Aux cable; there are 8 pairs in total.
PCI_BUSXY_WIDTH	Bandwidth of the PCIe bus	0 = inactive 1 = x1 2 = x2 3 = x4 4 = x8 5 = x16
PCI_BUSXY_SPEED	Link speed of the PCIe bus	0 = Gen1 1 = Gen2 2 = Gen3 3 = Gen4 4 = Gen5 5 = Gen6
PCI_BUSXY_HIERARCHY_TYPE	Hierarchy type of the PCIe bus	0 = PCIE_ENDPOINT 1 = PCIE_EXTERNAL_HOST_SWITCH 2 = PCIE_INTERNAL_HOST_SWITCH

Configuring the adapter

Now that the tools are installed, use the following procedure to configure the adapter to have a PCIe Gen5 x32 host interface.

1. To begin, run the following command to create the devices

```
# mst start
```

```
[root@HydraSIT-1 ~]# mst start
Starting MST (Mellanox Software Tools) driver set
Loading MST PCI module - Success
Loading MST PCI configuration module - Success
Create devices
Unloading MST PCI module (unused) - Success
```

Figure 4. mst start command

2. Check which adapters are installed on the server by running the following command:

```
# mst status
```

```
[root@HydraSIT-1 ~]# mst status
MST modules:
------------
    MST PCI module is not loaded
    MST PCI configuration module loaded

MST devices:
------------
/dev/mst/mt4131_pciconf0         - PCI configuration cycles access.
                                   domain:bus:dev.fn=0000:4c:00.0 addr.reg=88 data.reg=92 cr_bar.gw_offset=-1
                                   Chip revision is: 00
```

Figure 5. mst status command

In this output, you can see that there is only one adapter card on one PCI slot.

3. Run lspci to verify the card types

```
# lspci | grep Mellanox
```

```
[root@HydraSIT-1 ~]# lspci | grep Mellanox
4c:00.0 Ethernet controller: Mellanox Technologies CX8 Family [ConnectX-8]
4c:00.1 Ethernet controller: Mellanox Technologies CX8 Family [ConnectX-8]
```

Figure 6. lspci check

4. Run the following lspci command using the device number from Figure 4, to determine the current mode that the adapter is running in.

```
# lspci -s 4c:00.0 -vv
```

Version of lspci: Only recent versions of the pciutils package (which includes lspci) can recognize the rate of PCIe Gen6. Versions of lspci 3.10.0 and above have been tested and confirmed to support PCIe Gen6, however the inbox version of pciutils does not support Gen6.

Download the latest pciutils from https://github.com/pciutils/pciutils. If your OS is connected to the remote YUM repository, it can be upgraded online.

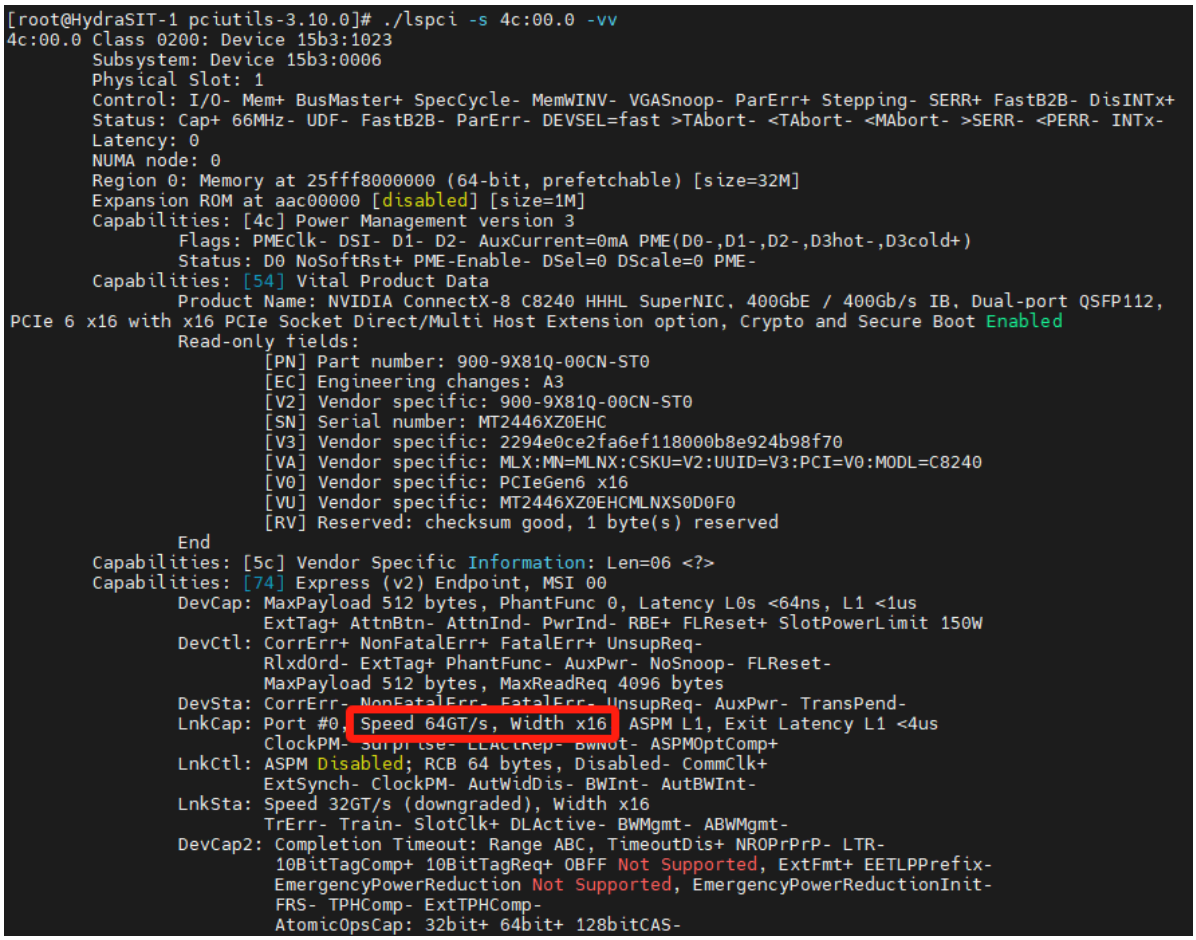

Figure 7. PCI capabilities of adapter

The values highlighted **Speed 64GT/s, Width x16** indicates that the adapter is currently configured to operating in the Gen6 x16 mode.

The output of the mst status command (Figure 5) includes the name of configuration node for the CX8 adapter. In our example, it is `/dev/mst/mt4131_pciconf0`. This value is used in the commands in the next steps.

Command syntax: For details about the `mlxconfig` commands used in the following steps, see the Parameters of the mlxconfig command section.

5. Set the adapter adjust the timeout of PCIe credit using the following command, as shown in the following figure.

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCIE_CREDIT_TOKEN_TIMEOUT=0
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCIE_CREDIT_TOKEN_TIMEOUT=0

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC: 400GbE (default mode) / 400Gb/s IB: Dual-port
QSFP112; PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secure Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCIE_CREDIT_TOKEN_TIMEOUT                   0               0

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 8. Set PCIE_CREDIT_TOKEN_TIMEOUT=0

6. Set the CX8 adapter to the PCIe bandwidth of x16 using the following command, as shown in the following figure

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS00_WIDTH=5
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS00_WIDTH=5

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC; 400GbE (default mode) / 400Gb/s IB; Dual-port QSFP112;
PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secure Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCI_BUS00_WIDTH                             PCI_INACTIVE(0)     PCI_X16(5)

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 9. Set PCI_BUS00_WIDTH=5

7. Set the CX8 adapter to the PCIe Gen5 using the following command, as shown in the following figure.

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS00_SPEED=4
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS00_SPEED=4

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC: 400GbE (default mode) / 400Gb/s IB; Dual-port QSFP112;
PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secure Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCI_BUS00_SPEED                             PCI_GEN_1(0)        PCI_GEN_5(4)

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 10. Set PCI_BUS00_SPEED=4

8. Set the CX8 adapter's hierarchy as PCIe endpoint using the following command, as shown in the following figure.

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS00_HIERARCHY_TYPE=0
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS00_HIERARCHY_TYPE=0

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC; 400GbE (default mode) / 400Gb/s IB; Dual-port QSFP112;
PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secu re Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCI_BUS00_HIERARCHY_TYPE                  PCIE_ENDPOINT(0)     PCIE_ENDPOINT(0)

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 11. Set PCI_BUS00_HIERARCHY_TYPE=0

9. Set the auxiliary cable connection to the PCIe bandwidth of x16 using the following command, as shown in the following figure.

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS10_WIDTH=5
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS10_WIDTH=5

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC; 400GbE (default mode) / 400Gb/s IB; Dual-port QSFP112;
PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secu re Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCI_BUS10_WIDTH                           PCI_INACTIVE(0)      PCI_X16(5)

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 12. Set PCI_BUS10_WIDTH=5

10. Set the auxiliary cable connection to the PCIe Gen5 using the following command, as shown in the following figure.

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS10_SPEED=4
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS10_SPEED=4

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC; 400GbE (default mode) / 400Gb/s IB; Dual-port QSFP112;
PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secure Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCI_BUS10_SPEED                           PCI_GEN_1(0)         PCI_GEN_5(4)

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 13. Set PCI_BUS10_SPEED=4

11. Set the auxiliary cable connection's hierarchy as PCIe endpoint using the following command, as shown in the following figure.

```
# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS10_HIERARCHY_TYPE=0
```

```
[root@HydraSIT-1 ~]# mlxconfig -d /dev/mst/mt4131_pciconf0 -y set PCI_BUS10_HIERARCHY_TYPE=0

Device #1:
----------

Device type:        ConnectX8
Name:               900-9X81Q-00CN-ST0_Ax
Description:        NVIDIA ConnectX-8 C8240 HHHL SuperNIC; 400GbE (default mode) / 400Gb/s IB; Dual-port QSFP112;
PCIe 6 x16 with x16 PCIe Socket Direct / Multi Host Extension option; Crypto Enabled; Secure Boot Enabled
Device:             /dev/mst/mt4131_pciconf0

Configurations:                                     Next Boot       New
        PCI_BUS10_HIERARCHY_TYPE                  PCIE_ENDPOINT(0)     PCIE_ENDPOINT(0)

 Apply new Configuration? (y/n) [n] : y
Applying... Done!
-I- Please reboot machine to load new configurations.
```

Figure 14. Set PCI_BUS10_HIERARCHY_TYPE=0

12. Restart the server to ensure the commands take effect.

Verifying the configuration

To verify that the adapter is now configured for PCIe Gen5 x32 mode, perform the following steps:

1. Run the mst start and mst status commands. You should see in the out that there is now an additional device, as highlighted in the following figure.

```
mst start
mst status
```

```
[root@HydraSIT-1 ~]# mst start
Starting MST (Mellanox Software Tools) driver set
Loading MST PCI module - Success
Loading MST PCI configuration module - Success
Create devices
Unloading MST PCI module (unused) - Success
[root@HydraSIT-1 ~]# mst status
MST modules:
------------
    MST PCI module is not loaded
    MST PCI configuration module loaded

MST devices:
------------
/dev/mst/mt4131_pciconf0         - PCI configuration cycles access.
                                   domain:bus:dev.fn=0000:4c:00.0 addr.reg=88 data.reg=92 cr_bar.gw_offset=-1
                                   Chip revision is: 00
/dev/mst/mt4131_pciconf1         - PCI configuration cycles access.
                                   domain:bus:dev.fn=0000:95:00.0 addr.reg=88 data.reg=92 cr_bar.gw_offset=-1
                                   Chip revision is: 00
```

Figure 15. Verify the existence of the new device

2. Executing the lspci command reveals an additional device, as shown in the following figure.

```
# lspci | grep Mellanox
```

```
[root@HydraSIT-1 ~]# lspci | grep Mellanox
4c:00.0 Ethernet controller: Mellanox Technologies CX8 Family [ConnectX-8]
4c:00.1 Ethernet controller: Mellanox Technologies CX8 Family [ConnectX-8]
95:00.0 Ethernet controller: Mellanox Technologies CX8 Family [ConnectX-8]
95:00.1 Ethernet controller: Mellanox Technologies CX8 Family [ConnectX-8]
```

Figure 16. lspci check after reboot

3. Verify that the ConnectX-8 adapter, which becomes Gen5 x16, as shown in the following figure. The value of the `-s` parameter comes from the output in Step 1. Verify that the output says "Speed 32GT/s, Width x16" as highlighted.

```
# lspci -s 4c:00.0 -vv
```

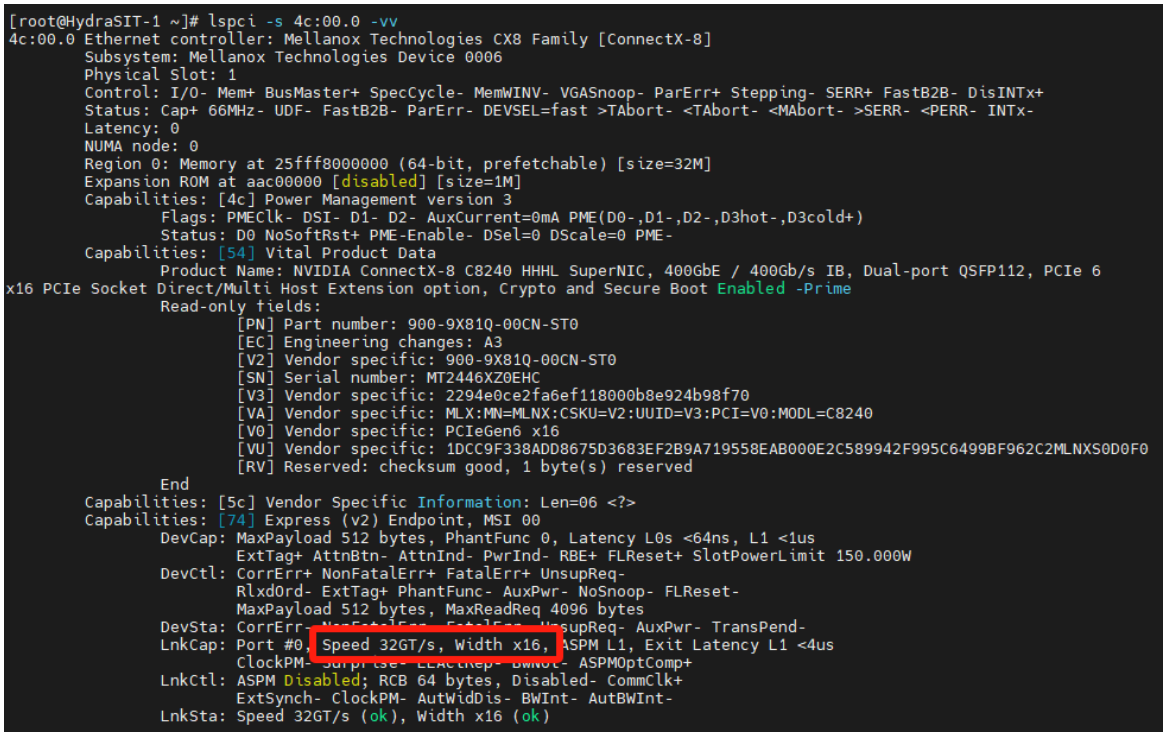

Figure 17. PCI property detection of the CX8 adapter card

4. Retrieve the PCIe properties of Auxiliary Cable, which should now be Gen5 x16, as shown in the following figure. The value of the `-s` parameter comes from the output in Step 1. Verify that the output says "Speed 32GT/s, Width x16" as highlighted.

```
# lspci -s 95:00.0 -vv
```

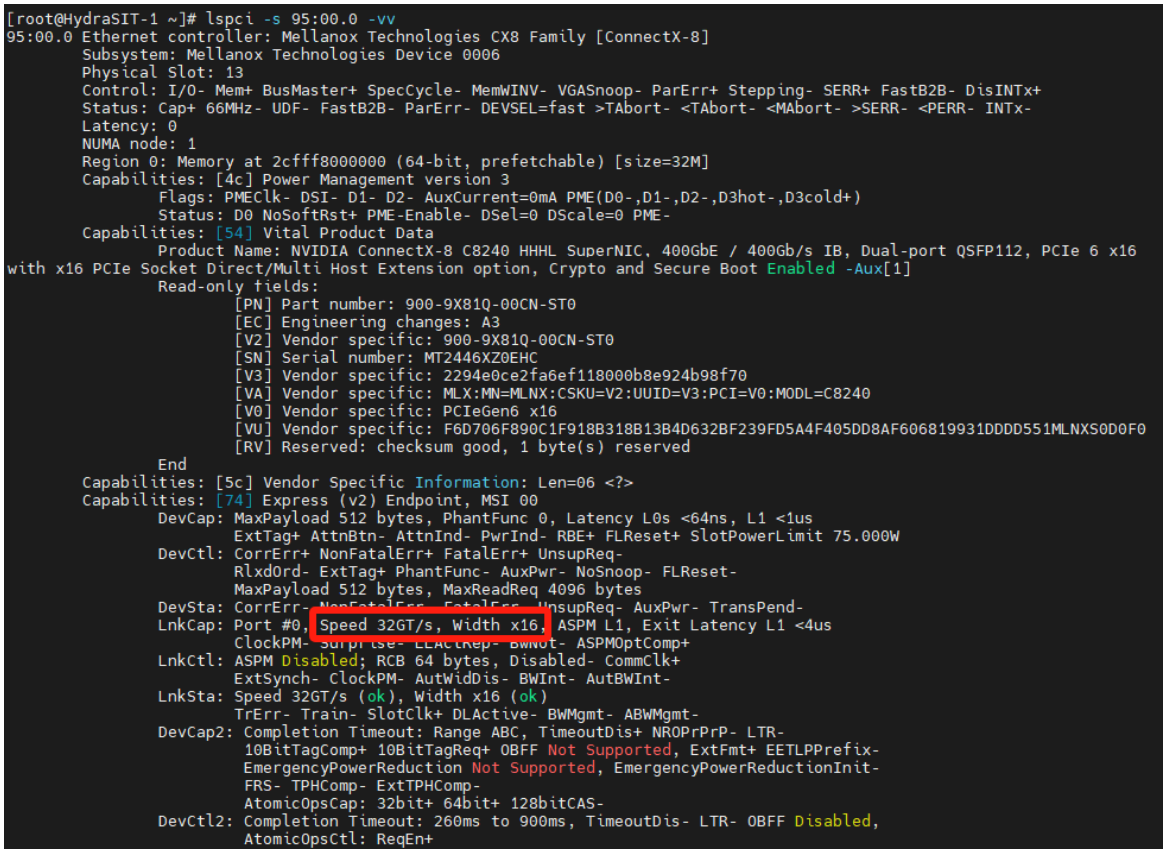

Figure 18. PCI property detection of the auxiliary cable

Resources

For more information, see the following web pages:

- Product Guide: ThinkSystem NVIDIA ConnectX-8 8180 800Gbs OSFP PCIe Gen6 x16 Adapter
  https://lenovopress.lenovo.com/lp2163-thinksystem-nvidia-connectx-8-8180-800gbs-osfp-pcie-gen6-x16-adapter
- Product Guide: ThinkSystem NVIDIA ConnectX-8 8240 400Gbs QSFP112 2-Port PCIe Gen6 x16 Adapter
  https://lenovopress.lenovo.com/lp2164-thinksystem-nvidia-connectx-8-8240-400gbs-qsfp112-2-port-pcie-gen6-x16-adapter
- NVIDIA User Manuals
  https://docs.nvidia.com/networking/index.html

Author

Liyong Duan is a Lenovo Senior Engineer responsible for NVIDIA Ethernet adapters in Lenovo ThinkSystem servers. Liyong Duan has worked in the IT industry since 2020, and is currently based in Tianjinn, China.

Related product families

Product families related to this document are the following:

- Ethernet Adapters

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
8001 Development Drive
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document, LP2227, was created or updated on May 30, 2025.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP2227
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP2227.

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at https://www.lenovo.com/us/en/legal/copytrade/.

The following terms are trademarks of Lenovo in the United States, other countries, or both:
Lenovo®
ThinkSystem®

Other company, product, or service names may be trademarks or service marks of others.