NVIDIA BlueField Reset and Reboot Procedures

Table of contents

BlueField System Reboot

This section describes the necessary operations to load new NIC firmware, following NVIDIA® BlueField® NIC firmware update. This procedure deprecates the need for full server power cycle.

The following steps are executed in the BlueField OS:

1. Issue a query command to ascertain whether BlueField system reboot is supported by your environment:

```
mlxfwreset -d 03:00.0 q
```

   If the output includes the following lines, proceed to step 2:

```
3: Driver restart and PCI reset
    -Supported (default)
...
1: Driver is the owner
-Supported (default)
```

> **Note**
>
> If it says `Not Supported` instead, then proceed to the instructions under section "BlueField System-level Reset".

2. Issue a BlueField system reboot:

```
mlxfwreset -d 03:00.0 -y -l 3 --sync 1 r
```

BlueField System-level Reset

This section describes the way to perform system-level reset (SLR) which is necessary for firmware configuration changes to take effect.

- SLR for BlueField running in DPU mode
- SLR for BlueField running in NIC mode
- SLR for BlueField running in DPU mode on hosts with separate power control (special use case)

System-level Reset for BlueField in DPU Mode

The following is the high-level flow of the procedure:

1. Graceful shutdown of BlueField Arm cores.
2. Query BlueField state to affirm shutdown reached.

   > **Info**
   >
   > In systems with multiple BlueField networking platforms, repeat steps 1 and 2 for all devices before proceeding.

3. Warm reboot the server.

Step by step process:

> **Info**
>
> Some of the following steps can be performed using different methods, depending on resource availability and support in the user's environment.

1. Graceful shutdown of BlueField Arm cores.

> **Info**
>
> This operation is expected to finish within 15 seconds.

Possible methods:

- From the BlueField OS:

```
shutdown -h now
```

  Or:

```
mlxfwreset -d /dev/mst/mt*pciconf0 -l 1 -t 4 --sync 0 r
```

- From the host OS:

> **Info**
>
> Not relevant when the BlueField is operating in Zero-Trust Mode.

```
mlxfwreset -d <mst-device> -l 1 -t 4 r
```

- Using the BlueField BMC:

```
ipmitool -C 17 -I lanplus -H <bmc_ip> -U root -P
<password> power soft
```

Or using Redfish (BlueField-3 and above):

```
curl -k -u root:<password> -H "Content-Type: application/json" -X
POST https://<bmc_ip>/redfish/v1/Systems/Bluefield/Actions/ComputerSystem.Reset -d
'{"ResetType": "GracefulShutdown"}'
```

2. Query BlueField state.

   Possible methods:

   - From the host OS:

     Info

     Not relevant when the BlueField is operating in Zero-Trust Mode.

     ```
     echo DISPLAY_LEVEL 2 > /dev/rshim0/misc
     cat /dev/rshim0/misc
     ```

     Expected output:

     ```
     INFO[BL31]: System Off
     ```

   - Utilizing the BlueField BMC:

```
ipmitool -C 17 -I lanplus -H <bmc_ip> -U root -P
<password> raw 0x32 0xA3
```

Expected output: `06`.

3. Warm reboot the server f rom the host OS:

```
mlxfwreset -d <mst-device> -l 4 r
```

> **Note**
>
> If multiple BlueField devices are present in the host, this command must run only once. In this case, the MST device can be of any of the BlueFields for which the reset is necessary and participated in step 1.

Or:

```
reboot
```

> **Note**
>
> For external hosts which do not toggle PERST# in their standard reboot command, use the `mlxfwreset` option.

System-level Reset for BlueField in NIC Mode

Perform warm reboot of the host OS:

```
mlxfwreset -d <mst-device> -l 4 r
```

Or:

```
reboot
```

> **Note**
>
> For external hosts which do not toggle PERST# in their standard reboot command, use the `mlxfwreset` option.

System-level Reset for Host with Separate Power Control

This procedure is a special use case relevant only to host platforms with separate power control for the PCIe slot and CPUs, in which the BlueField (running in DPU mode) is provided power while host OS/CPUs may be in shutdown or similar standby state (this allows the BlueField device to be operational while the host CPU is in shutdown/standby state).

The following is the high-level flow of the procedure:

1. Graceful shutdown of host OS or similar CPU standby.
2. Graceful shutdown of BlueField Arm cores.
3. Query BlueField state to affirm shutdown reached.

4. Full BlueField Reset

5. Query BlueField state to affirm operational state reached

> **Info**
>
> In systems with multiple BlueField networking platforms, repeat steps 1 through 5 for all devices before proceeding.

6. Power on the server.

Step by step process:

> **Info**
>
> Some of the following steps can be performed using different methods, depending on resource availability and support in the user's environment.

1. Graceful shutdown of host OS by any means preferable.

2. Graceful shutdown of BlueField Arm cores.

> **Info**
>
> This step normally takes up to 15 seconds to complete.

- From the BlueField OS:

```
shutdown -h now
```

- Utilizing the BlueField BMC:
  - Using IPMI:

```
ipmitool -C 17 -I lanplus -H <bmc_ip> -U root -P
<password> power soft
```

  - Using Redfish (for BlueField-3 and above):

```
curl -k -u root:<password> -H "Content-Type: application/json" -X
POST
https://<bmc_ip>/redfish/v1/Systems/Bluefield/Actions/ComputerSystem.Reset -d
'{"ResetType": "GracefulShutdown"}'
```

3. Query the BlueField's state utilizing the BlueField BMC:

```
ipmitool -C 17 -I lanplus -H <bmc_ip> -U root -P <password>
raw 0x32 0xA3
```

   Expected output: `06`.

4. Perform BlueField hard reset utilizing the BlueField BMC:

> **Info**
>
> This step takes up to 2 minutes to complete .

- Using IPMI:

```
ipmitool -C 17 -I lanplus -H <bmc_ip> -U root -P
<password> power cycle
```

- Using Redfish (for BlueField-3 and above):

```
curl -k -u root:<password> -H "Content-Type: application/json" -X
POST https://<bmc_ip>/redfish/v1/Systems/Bluefield/Actions/ComputerSystem.Reset -d
'{"ResetType" : "PowerCycle"}'
```

5. Query BlueField operational state u tilizing the BlueField BMC :

> **Info**
>
> At this point, the BlueField is expected to b e operational .

```
ipmitool -C 17 -I lanplus -H <bmc_ip> -U root -P <password>
raw 0x32 0xA3
```

Expected output: `05`.

6. Power on/boot up the host OS.

BlueField DPU Reset Using a BMC Platform

The BlueField DPU can also be reset from a BMC platform using NC-SI command over I2C. This option is more common in DPU BMC absence, when the BlueField DPU is running in

NIC mode or when it is used as a controller.

The reset is performed using the *Reset BlueField-3 DPU (Command=0x12, Parameter=0xB)*, which allows a BMC platform to reset the NVIDIA BlueField-3 DPU device. This command is only applicable to BlueField-3 devices.

The Reset BlueField-3 DPU command is addressed to the package only. When the internal reset is complete, the BMC platform should reconfigure the device.

> **Info**
>
> The *Reset BlueField-3 DPU* command is supported on BlueField-2 and later devices.

Reset BlueField-3 DPU Format

Bytes/Bits	31:24	23:16	15:8	7:0	
0:15	NC-SI Header (OEM Command)				
16:19	NVIDIA Manufacture ID (IANA) = 0x8119				
20:23	Command rev=0x00	MLNX Cmd ID=0x12	Parameter=0x0B	NICR	Mode
24:27	Checksum 31:0				

The parameter descriptions for Reset BlueField-3 DPU command are provided below.

Reset BlueField-3 DPU Parameters

Field	Description
NICR	• 0 - NIC does not reset. Only the embedded CPU will reset. • 1 - Reset the embedded CPU and the NIC
Mode	This field defines the type of conditions to use before performing the internal reset • 0 - The internal reset will start after sending the response to this command • 1 - The internal reset will start only when all the hosts asserts their PERST# signals low

Field	Description
	• 2 - The internal reset will start only when all the hosts disabled their PCIe links. This may or may not include assertion of their respective PERST# signals low. • Other - Reserved

Reset BlueField-3 DPU Response

The ConnectX adapter responds to a Reset BlueField-3 DPU command when the package ID matches, and with no checksum error.

Reset BlueField-3 DPU Response Format

Bytes/Bits	31:24	23:16	15:8	7:0	
0:15	NC-SI Header (OEM Command)				
16:19	Response Code		Reason Code		
20:23	NVIDIA Manufacture ID (IANA) = 0x8119				
24:27	Command rev=0x00	MLNX Cmd ID=0x12	Parameter=0x0B	NICR	Mode
28:31	Checksum 31:0				

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.

Trademarks

NVIDIA and the NVIDIA logo are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.
