NVIDIA BlueField DPU Time and Scheduling Deployment Guide for Rivermax (Windows)

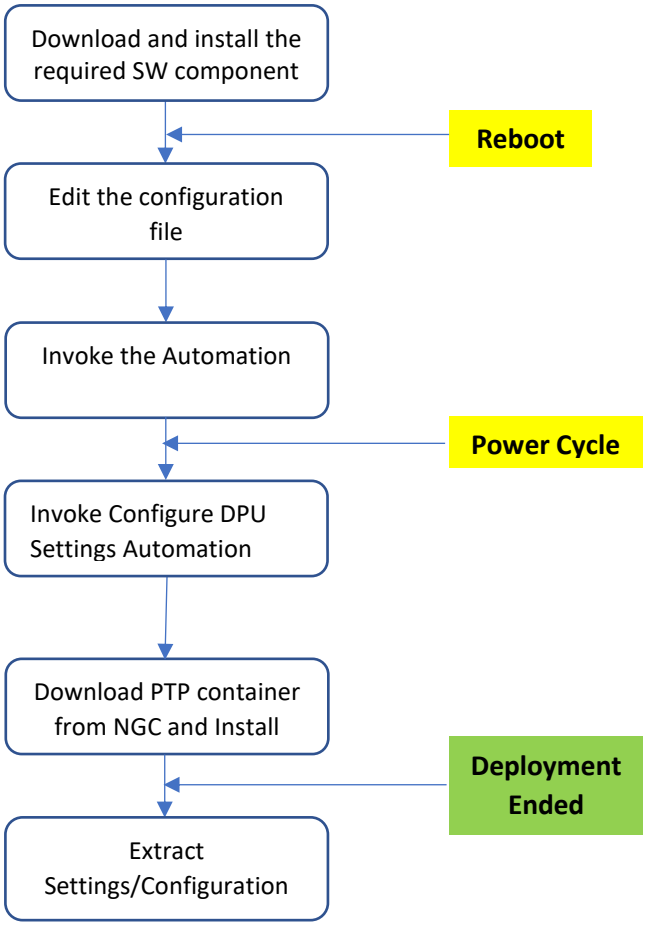
Version 1.09
Date 30-March-22

Table of Contents

1 DPU Deployment Steps and Flow Diagram 3
2 Physical Setup 4
3 Software Components 4
4 NVIDIA BlueField-2 Cables connection 4
5 Prerequisites 5
6 PTP Pre-Requisite 6
7 DPU Setup – Start automation scripts 8
8 DPU Settings 8
9 DPU PTP Service Container 9
10 Basic Troubleshooting 10
11 Recommendations 13

1 DPU Deployment Steps and Flow Diagram

1. Download and install a few SW component on the Windows server (WinOF-2, Python, PuTTy and the BFB image)
2. Editing a configuration file with information like BlueField-2 IP, password, and gateway
3. Run Automation from windows server command line – [Power Cycle is required]
4. Configure DPU settings
5. PTP Container deployment from NGX using putty (opens automatically)
6. Extract settings – this is done automatically for future debug



2 Physical Setup

3 Software Components

The following table presents the software components used to deploy NVIDIA® BlueField® DPU:

Component	Download Link	Version
BlueField BFB (BlueField-2 Ubuntu Server **20.04 - BlueField OS image)**	Direct Link Page Link	**Note: due to an issue with the latest BFB on the web (1.2.1) – BFB 1.2.0 is located in the BFB directory as part of the provided ZIP file**
PuTTY	Link	0.76 or later
Python	Link	3.10.1 or later
MLNX_WinOF2	Link	2.80 or later
BlueField Firmware		24.32.1010 [Part of the BFB image and set by WinOF-2]

For latest software versions -
https://docs.nvidia.com/networking/display/BlueFieldDPUOSv380/Supported+Platforms+and+Interoperability

4 NVIDIA BlueField-2 Cables connection

Connect 2 types of cables:

- miniUSB Cable is not needed – can be connected for terminal usage
- Ethernet cable - used for **DPU management**, connected to the **OOB port**. DPU is configured by default to acquire IP via DHCP - make sure to configure your DHCP server with the DPU OOB MAC
- **Port 0** - 100GbE cable connected to the high-speed switch

5 Prerequisites

1. Unzip the ZIP file to this path on the Windows server: **C:\DPU**

   Note: the path is used hardcoded in the scripts

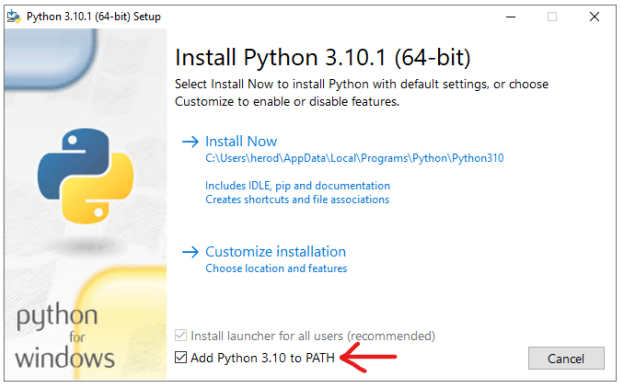
   The ZIP file contains automated scripts to install and configure the DPU.

   The Zip file contains:

   - DPU_install.batch – main script that start the process.
   - automation.batch – contains the commands for the BFB and FW installations.
   - configure_dpu_setting – contains the commands and scripts for the DPU configurations.
   - configuration_file.txt – file for the user to insert his info.
   - create_container.sh – script that run on the DPU and create the container.
   - remote_commands.txt – using this file running commands remotely on the DPU.
   - reset_password.py – script that change the password of the system to the user password.
   - set_ip_and_container.sh – set IP for the PTP interface, install needed package to run the container.
   - set_real_time_clock.sh – enable real time clock.
   - setting_status.sh - print the current DPU status.
   - Read_DPU_Configuration.batch - main script to get the current DPU status.
   - set_persistent_ip_to_p0.sh – set persistent IP and gateway to the PTP interface.
   - FW_Update.batch – force update the DPU FW to the latest FW version.
   - remote_commands_dns_fix.txt – using this file to run commands on the DPU to fix the DNS.
   - get_fw_info.batch – get the current device FW version and status.
   - config_after_reboot.sh – set the PTP interface Tx Timestamp to on after each reboot.
   - rc.local – will copied to the Bluefield /etc/ and run the "config_after_reboot" script after each reboot.

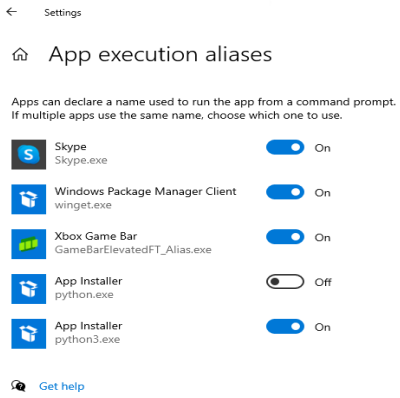
2. Download the software components from section 3
   1. Download the BFB image and place it in this folder **C:\DPU\BFB**
   2. **Note: due to an issue with the latest BFB on the web (1.2.1) – BFB 1.2.0 is located in the BFB directory as part of the provided ZIP file**
   3. Install PuTTY
   4. Install MLNX_WinOF2
      - For Rivermax installation please refer to the Rivermax installation guide
      - Rivermax version 1.9.x installation requiring a registry key - **command line must be used: MLNX_WinOF2-2_80_50000_All_x64.exe /v"MT_RIVERMAX=1 /l*vx C:\Users\\log.txt "**
   5. Install app ex

      Enable "add Python 3.10 to PATH"



After the Python 3 installation, type at the search bar "App execution aliases"

Make sure that only python3.exe is on (as the image below).



***Note: A Reboot is required after installation**

3. Edit the configuration_file at **C:\DPU\configuration_file.txt**
   - **BlueField-2_Managment_IP**: BlueField-2 DPU OOB Management IP
   - **BlueField-2_Managment_Gateway**: BlueField-2 DPU OOB Management Gateway
   - **BlueField-2_PTP_IP_Interface_0**: DPU arm PTP interface IP address and Subnet Mask
   - **BlueField-2_PTP_IP_Interface_0_Gateway**: DPU arm PTP interface Gateway
   - **BlueField-2_Username**: BlueField-2 current username (can be "root")
   - **BlueField-2_Password**: BlueField-2 current password
   - **BlueField-2_Root_Password**: BlueField-2 root current password
   - **Force_BFB_image_update(y/n):** y – force install BFB image (default), n – the BFB image will install only if the current BFB version is different from the BFB image version at C:\DPU\BFB\

***NOTES:**

- It is mandatory to fill in all the fields
- If you do not know the root password, BFB burn required!
- After the installation of the BFB the system requires a password change.

BlueField-2_Password requirements:

- Minimum password length of **8 characters**
- Use at least one digit
- Use at least one Special character
- Don't use the Special character: **!** (Exclamation mark)

6 PTP Pre-Requisite

Configure PTP on the switch side, while ensuring that the Bluefield IP and PTP are using the same subnet.

More information is available as part of the PTP best practice document.

*Make sure you have access permissions to container on NVIDIA NGC

7 DPU Setup – Start automation scripts

Open a Command line as an administrator and run the script file **C:\DPU\DPU_Install.bat**

The installation will take ~15 minutes, at the end the automation will perform a FW upgrade.

After FW upgrade a power cycle to the server is needed – Manually

8 DPU Settings

After the BFB installation and FW upgrade it's required to configure a few settings.

Open a Command line as an administrator and run the script file **C:\DPU\configure_dpu_setting.bat**

Once completed a PuTTY window will be opened to set the PTP container [next section]

Another window will open and all the DPU settings will be printed for debug/tracking.

9 DPU PTP Service Container

1. Open a new command line window: putty -ssh root@BlueFied_2_Managment_IP 22 -pw BlueField-2_Password
2. In the PuTTY window Run the following command to start the PTP Container setup:

   # ./create_container.sh

3. Insert the interface name for PTP usage (i.e. p0):

   # p0
4. Insert Username and Password, see below:

   Note: The script will download a container from the NVIDIA NGC catalog and require providing an API key to log into the NGC registry:

   Username: $oauthtoken
   Password: <Your_NGC_API_Key>

   For private container (early access program) see NGC Private Registry User Guide -> section "Generating Your NGC API Key"

   *Make sure you have access permissions to container on NVIDIA Rivermax NGC

5. Verify the container is up and running:

   # docker ps

```
CONTAINER ID    IMAGE
a8df01c70d6e    nvcr.io/ea-rivermax/dpu-ptp-service/dpu-ptp-serice:21.03-
v1
COMMAND         CREATED        STATUS       PORTS     NAMES
"./start.sh"    45 hours ago    Up 45 hours           ptp-service-container
```

6. Check /var/log/ptp4l.log file for PTP lock:

   # tail -f /var/log/ptp4l.log

```
ptp4l[326.286]: rms    5 max    9 freq   +945 +/-   4 delay   109 +/-   1
ptp4l[327.286]: rms    4 max   10 freq   +944 +/-   5 delay   109 +/-   0
ptp4l[328.286]: rms    4 max    8 freq   +946 +/-   5 delay   109 +/-   1
ptp4l[329.286]: rms    5 max    8 freq   +942 +/-   6 delay   108 +/-   0
ptp4l[330.286]: rms    5 max   10 freq   +944 +/-   7 delay   108 +/-   1
ptp4l[331.286]: rms    3 max    6 freq   +945 +/-   5 delay   110 +/-   0
```

10 Basic Troubleshooting

BFB installation Timeout

Issue - Timeout occurs when trying to install BFB-write error: Connection timed out

The timeout may be due to a USB connection issue; Try these steps to resolve the problem:

- On the server: remove MLNX_WinOF2, reboot and reinstall (make sure latest version is used)
- Use a different USB port
- If you have another USB cable, replace it
- Try to software reset the DPU using the following command on the Windows server:

  RshimCmd -RestartSmartNic 1 -BusNum 3

  Note: The BusNum should be taken from ***rshimcmd -enumdevices***

  - Guidelines for RSHIM support https://docs.nvidia.com/networking/display/winof2v280/RShim+Drivers+and+Usage
- Try to power cycle the server hosting the DPU
- **Error:** 'RshimCmd' is not recognized as an internal or external command
  - Make sure to restart the server after WinOF-2 installation
- Ensure that the FW version is correct – a force update might be required – use mlxup software - https://www.mellanox.com/support/firmware/mlxup-mft
- In extreme use cases it might be recommended to restore the DPU to default
  - mlxfwreset -d XX:XX.X -y reset
  - XX:XX.X = pci address

Extract Settings and Configurations

- At the end of the deployment the setting and configuration will be printed -
- Verify that you get the same “Setting Status” as below or extract is it using the Read_DPU_Configuration.bat
- Note: tx_port_ts need to be set on every reset – it will be part of the container in the next releases

```
DPU Automation Version:0.99
PCI Bus:3
-------- Setting Status --------------
Starting MST (Mellanox Software Tools) driver set
Loading MST PCI module - Success
[warn] mst_pciconf is already loaded, skipping
Create devices
Unloading MST PCI module (unused) - Success
Real Time:
        REAL_TIME_CLOCK_ENABLE              True(1)
Loopback:
        MPFS_MC_LOOPBACK_DISABLE_P1         True(1)
        MPFS_MC_LOOPBACK_DISABLE_P2         True(1)
Tx Timestamp:
tx_port_ts         : on
tx_port_ts         : on
OVS:
256b5be9-0765-4a2c-95b9-31f40ee7de49
    ovs_version: "2.15.1"
device mode:
        INTERNAL_CPU_MODEL                  SEPARATED_HOST(0)
privilege:
Host configurations
-------------------
level                          : PRIVILEGED

Port functions status:
----------------------
disable_rshim                  : FALSE
disable_tracer                 : FALSE
disable_port_owner             : FALSE
disable_counter_rd             : FALSE

interface IP:
p0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 1.1.2.3  netmask 255.255.255.0  broadcast 1.1.2.255
        inet6 fe80::bace:f6ff:fed1:3fb4  prefixlen 64  scopeid 0x20<link>
        ether b8:ce:f6:d1:3f:b4  txqueuelen 1000  (Ethernet)
        RX packets 59  bytes 3885 (3.8 KB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 28  bytes 2349 (2.3 KB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

```
BFB version:
DOCA_v1.2.1_BlueField_OS_Ubuntu_20.04-5.4.0-1023-bluefield-5.5-2.1.7.0-3.8.5.12027-1.signed-aarch64
FW version:
Querying Mellanox devices firmware ...

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

  Device Type:      BlueField2
  Part Number:      MBF2H516A-EENO_Ax_Bx
  Description:      BlueField-2 DPU 100GbE/EDR/HDR100 VPI Dual-Port QSFP56; PCIe Gen4 x16; Crypto Disabled; 16GB on-boar
d DDR; 1GbE OOB management; FHHL
  PSID:             MT_0000000705
  PCI Device Name:  /dev/mst/mt41686_pciconf0
  Base MAC:         b8cef6d13fb0
  Versions:         Current        Available
     FW             24.32.1010     N/A
     PXE            3.6.0502       N/A
     UEFI           14.25.0017     N/A
     UEFI Virtio blk   22.2.0010       N/A
     UEFI Virtio net   21.2.0010       N/A

  Status:           No matching image found
```

- ## No PTP after reboot
  - Make sure that P0 interface has a valid IP address (netplan has been applied during automation)

- ## Server of DPU is getting rebooted
  - If the BlueField-2 is installed on a server without proper cooling it might reach high temperature – it's advised to place an active fan on it to ensure constant air flow
- ## COM3 console can be used to monitor the deployment process
  - Launch PuTTY on port COM3 to view progress of the Ubuntu install on the DPU
- ## Program FW manually

  Open a Command line as administrator and run the script file **C:\DPU\FW_Update.bat**

  After FW upgrade a power cycle to the server is needed
- ## Not using latest Python
  - Make sure that during the Python installation you set the add Python to path
  - Make sure that no other Python version is in Windows Path environment variable

11 Recommendations

1. Tune the server that is hosting the DPU by Rivermax Windows tunning guide: https://community.mellanox.com/s/article/Rivermax-Windows-Performance-Tuning-Guide

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. Neither NVIDIA Corporation nor any of its direct or indirect subsidiaries (collectively: "NVIDIA") make any 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.

Trademarks

NVIDIA, the NVIDIA logo, and BlueField are trademarks and/or registered trademarks of NVIDIA Corporation and its affiliates in the U.S. and in other countries. Other company and product names may be trademarks of the respective companies with which they are associated. All other trademarks are property of their respective owners.

Copyright

© 2022 NVIDIA Corporation & Affiliates. All rights reserved.