Running NVIDIA DOCA Reference Applications Over Host

User Guide

MLNX-15-060490 _v1.1 | October 2021

Table of Contents

Chapter 1. Introduction

There might several reasons for running application on the host. For example, one might want to utilize a strong and high-resource host machine, or to start DOCA integration on the host before offloading it to the BlueField DPU.

The configuration in this document allows the entire application to run on the host's memory, while utilizing the HW accelerators on the BlueField (e.g., using RegEx the accelerator on BlueField using a daemon running on the DPU).

When virtual functions are enabled on the host, virtual function representors are visible on the Arm side which can be bridged to corresponding physical interface representors (e.g., the uplink representor and the host representor). This allows the application to only scan traffic being forwarded to the VFs as configured by the user and to behave as a simple "bump-on-the-wire" solution. DOCA installed on the host allows accessing the hardware capabilities of the BlueField DPU without comprising features such as the stateful table (SFT) which uses hardware offload and additional hardware steering elements embedded inside the eSwitch.

Chapter 2. Prerequisites

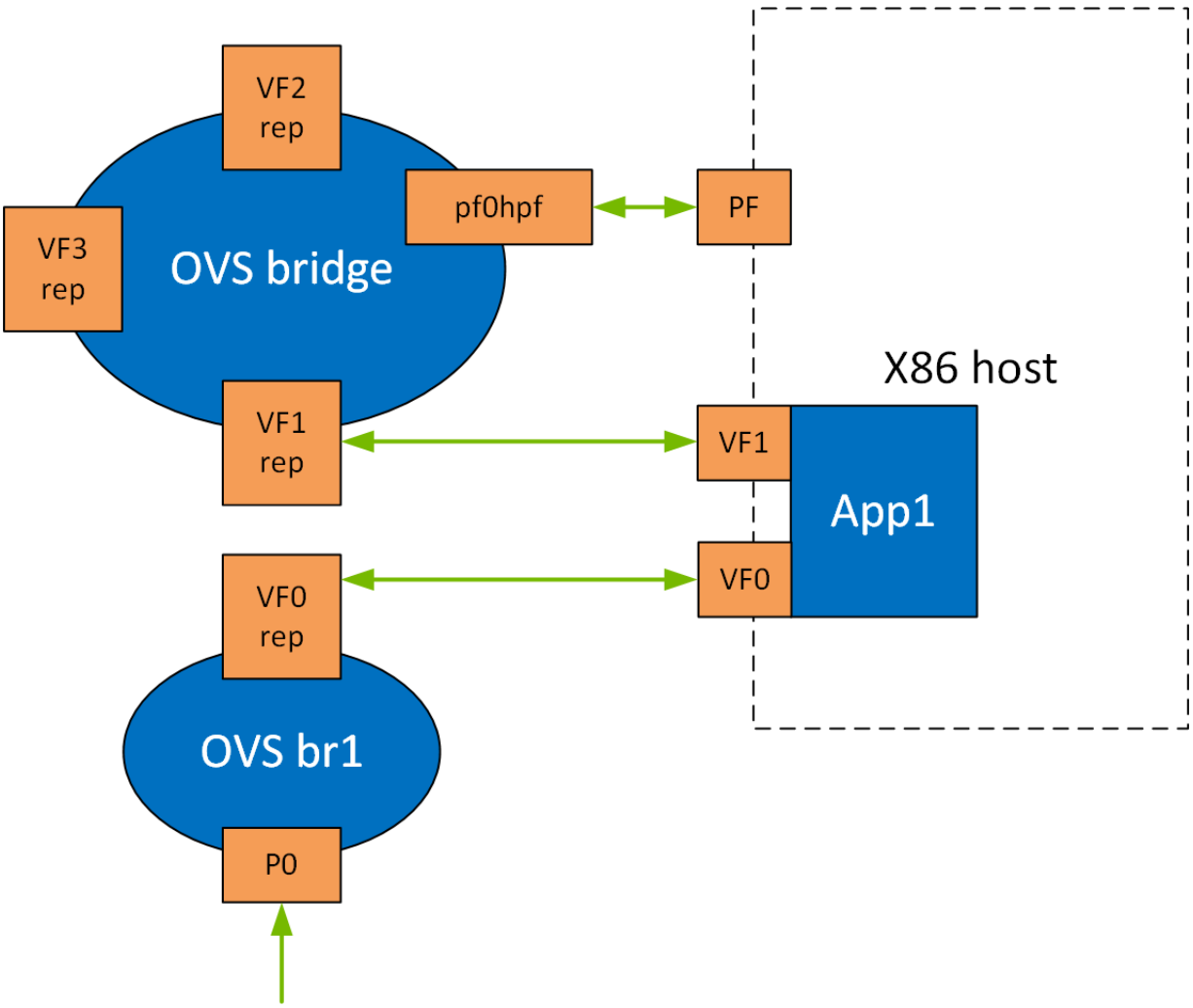
Running applications on the host and using the RegEx accelerator on the BlueField require enabling the RegEx engine.

To run all the reference applications over the host, it is required to install the host DOCA version of DPDK. For more information on host installation, please refer to DOCA Installation Guide.

Note: VFs must be configured as trusted for hardware jump action to work as intended. Run the following command on the Arm device:

```
mlxreg -d /dev/mst/mt41686_pciconf0 --reg_id 0xc007 --reg_len 0x40 --indexes
 "0x0.0:32=0x80000000" --yes --set "0x4.0:32=0x1"
```

To run the application over the host, the following topology is be used.



Configure the OVS on BlueField as follows:

```
    Bridge ovsbr1
        Port ovsbr1
            Interface ovsbr1
                type: internal
        Port pf0hpf
            Interface pf0hpf
        Port pf0vf1
            Interface pf0vf1
    Bridge vf_br
        Port p0
            Interface p0
        Port vf_br
            Interface vf_br
                type: internal
        Port pf0vf0
            Interface pf0vf0
```

Enabling VF over the host bridges the uplink connection (p0) to one end of the VF created (pf0vf0), and the other end (pf0vf1) to the host representors (pf0phf). On the host, the same application is run but 3 PCIe addresses are used:The first one for initiating the RegEx (must be done on interface 0), and the other 2 PCIe address to initialize the virtual function. Now, when traffic is received (e.g., from the uplink), the following flow occurs:

1. Traffic is received over p0.
2. Traffic is forwarded to pf0vf0.
3. Application "listens" to pf0vf0 and pf0vf1 and can, therefore, acquire the traffic from pf0vf0, inspect it, and forward to pf0vf1.
4. Traffic is forwarded from pf0vf1 to pf0hpf.

Chapter 3. Configuration Flow

1. Stop the driver on the host. Run:
   ```
   host$ sudo /etc/init.d/openibd stop
   ```
2. On the Arm, start the driver. Run:
   ```
   dpu$ sudo /etc/init.d/openibd start
   ```
3. On the Arm, enable RegEx. Run:
   ```
   dpu$ echo 1 > /sys/class/net/p0/smart_nic/pf/regex_en
   ```
4. On the Arm, add 200 huge pages. Run:
   ```
   dpu$ current_huge='cat /sys/kernel/mm/hugepages/hugepages-2048kB/nr_hugepages'
   dpu$ echo $((200 + current_huge)) > /sys/kernel/mm/hugepages/hugepages-2048kB/
   nr_hugepages
   ```
5. On the Arm, start mlx RegEx. Run:
   ```
   dpu$ systemctl start mlx-regex
   ```

   Note: If it has not been set before, the previous value of huge pages should be 2048 or higher (depending on the number of cores).
6. Verify that the service is running. Run:
   ```
   dpu$ systemctl status mlx-regex
   ```
7. The host can now run RegEx. Run:
   ```
   host$ sudo /etc/init.d/openibd start
   ```

Note: Running DPDK over the host requires configuring huge pages to be identical to the DPU configuration.

Running the application over the host also initializes all the cores of the device which is usually unnecessary and may even cause unforeseeable issues. It is recommended to limit the number of cores, especially when using an AMD-based system, to 16 cores using the `-c` flag when running DPDK.

The following is a CLI example for running a reference application over the host using VF:

```
./opt/mellanox/doca/example/**/bin/*executable* -a "pci address VF0" -a "pci address
 VF1" -c 0xff -- "application flags"
```

Note: The executable will fail if no correct `LD_LIBRARY_PATH` is set. To set `LD_LIBRARY_PATH`, execute the following.

- For Ubuntu:
  ```
  export LD_LIBRARY_PATH=/opt/mellanox/dpdk/lib/aarch64-linux-gnu/
  ```
- For CentOS:

```
export LD_LIBRARY_PATH=/opt/mellanox/dpdk/lib64
```

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 nor any of its direct or indirect subsidiaries and affiliates (collectively: "NVIDIA") make no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assume 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 Mellanox are trademarks and/or registered trademarks of Mellanox Technologies Ltd. and/or NVIDIA Corporation 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.

Copyright

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, CA 95051
http://www.nvidia.com