NVIDIA DOCA File Scan

Reference Application Guide

MLNX-15-060504 _v1.2 | January 2022

Table of Contents

Chapter 1. Introduction..... 1
Chapter 2. System Design..... 2
Chapter 3. Application Architecture..... 4
Chapter 4. Configuration Flow..... 5
Chapter 5. Running Application on BlueField..... 7
Chapter 6. Arg Parser DOCA Flags..... 9
Chapter 7. Running Application on Host..... 10
Chapter 8. References..... 11

Chapter 1. Introduction

BlueField DPU supports high-speed hardware RegEx acceleration using the RegEx engine. This allows accelerating different security and networking applications which require RegEx processing.

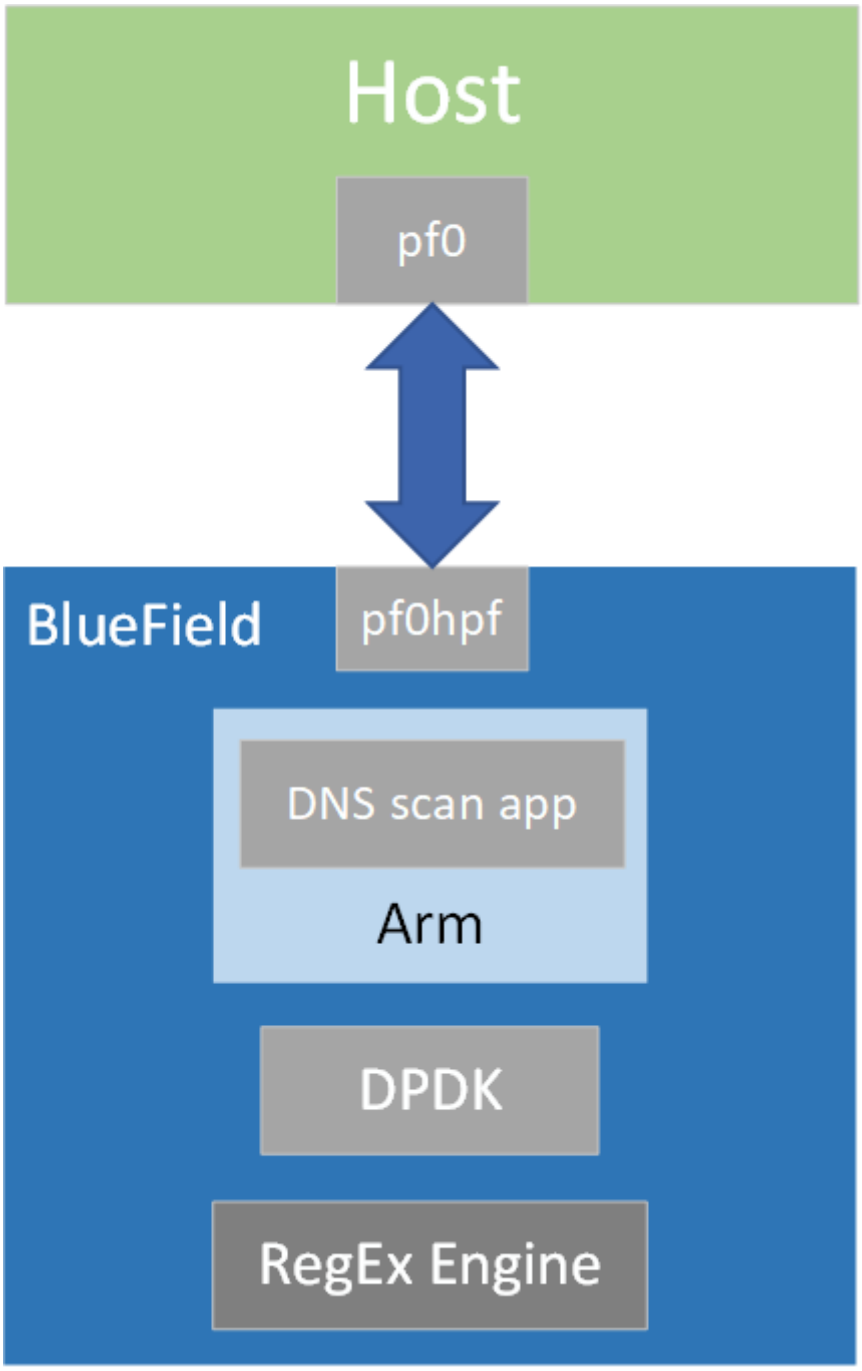
The file scan application uses the hardware RegEx engine to scan a file and returns whether there is a match or not. This application can be a base for many security verifications for sensitive data and personal information.

File scan logic includes:

- Compiling RegEx rules using Regex compiler
- Loading a compiled set of RegEx rules to the BlueField RegEx engine
- Sending file to the RegEx for scanning
- HW scanning file and returning a list of matches (e.g., matching rule, offset, length)

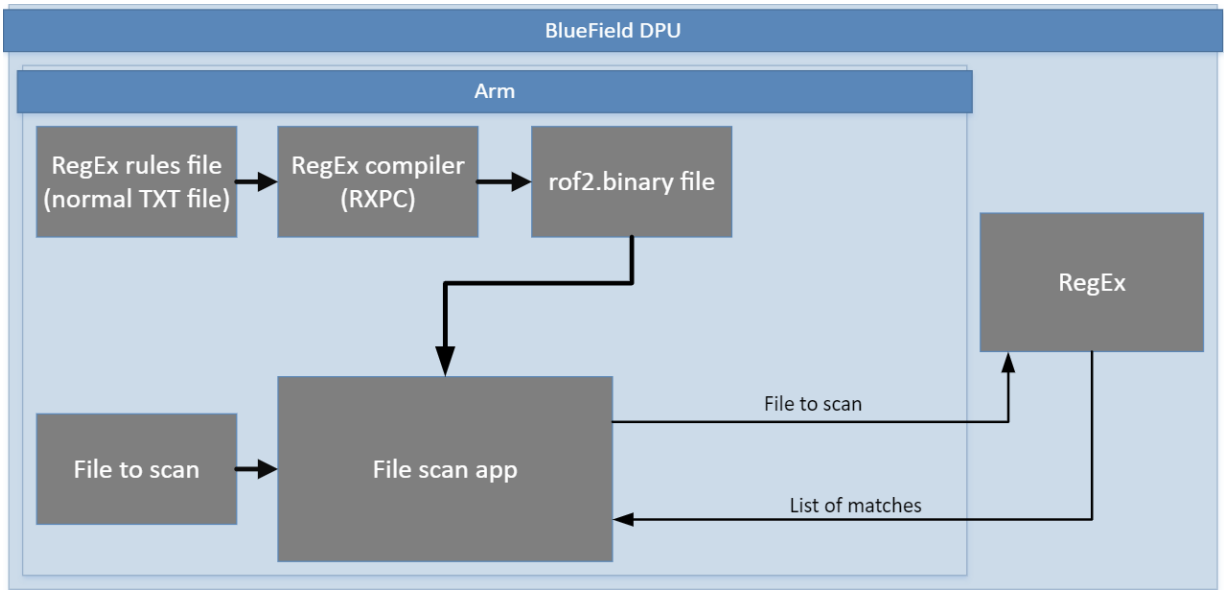
Chapter 2. System Design

File scan utilizes the RegEx engine which is HW accelerator on the BlueField.


Host
pf0
BlueField
pf0hpf
DNS scan app
Arm
DPDK
RegEx Engine

Chapter 3. Application Architecture

The file scan application runs on top of DPDK RegEx API to configure, send, and receive scans to and from the BlueField RegEx engine.



1. RegEx rules file is compiled into a rof2.binary file by the user.
2. The RegEx binary rules file is loaded into the RegEx engine.
3. The file to scan is sent to the RegEx triggering the engine to start matching

> **Note:** The current file scan version supports files up to 16KB in size.

4. The results are sent back to the application.

> **Note:** It is the application's responsibility to check for finished jobs. In this example, you Busy-Wait until the RegEx finishes scanning.

5. The matches are printed on the screen.

Chapter 4. Configuration Flow

1. Parse application argument.
   ```
   arg_parser_init();
   ```
   a). Initialize Arg Parser resources.
   b). Register DOCA general flags.
   ```
   register_file_scan_params();
   ```
   c). Register file scan application flags.
   ```
   arg_parser_start();
   ```
   d). Parsing DPDK flags and calling `rte_eal_init()` function.
   e). Parsing APP flags.
2. Initialize file scan resources.
   ```
   file_scan_init();
   ```
   a). Allocate buffer for the rules and data file.
3. Configure RegEx device.
   ```
   configure_regexdev();
   ```
   a). Configure RegEx engine with the predefined rules file.
   b). Configure RegEx queues.
4. Allocate mbuf pool.
   ```
   allocate_mbuf_pool();
   ```
   a). Allocate mbuf memory pool to hold the file buffer.
5. Load the file into the mbuf pool before enqueuing RegEx scan task.
   ```
   allocate_file_to_mempool();
   ```
6. Enqueue RegEx scan task.
   ```
   regex_enqueue_task();
   ```
   a). Send the file to the RegEx for scanning.
7. Dequeue RegEx scan task.
   ```
   rte_regexdev_dequeue_task();
   ```
   a). Busy-Wait until the file has been scanned.
8. Report results.
   ```
   report_results();
   ```
   a). Print the scan results to the standard output.
9. Destroy the file scan app.
   ```
   file_scan_destroy();
   ```

a). Free allocated buffers.
b). Free memory pool.

Chapter 5. Running Application on BlueField

1. Refer to the DOCA Installation Guide for details on how to install#BlueField#related software.
2. The application binary is located under `/opt/mellanox/doca/examples/file_scan/bin/doca_file_scan`.
3. To rebuild the application:

   a). Run:
   ```
   cd /opt/mellanox/doca/examples/file_scan/src
   meson /tmp/build
   ninja -C /tmp/build
   ```
   `doca_file_scan` is created under `/tmp/build`.

   b). The build process depends on the `PKG_CONFIG_PATH` environment variable to locate the DPDK libraries. If the variable has been accidently corrupted, and the build fails, run the following:

   - For Ubuntu:
     ```
     export PKG_CONFIG_PATH=$PKG_CONFIG_PATH:/opt/mellanox/dpdk/lib/aarch64-linux-gnu/pkgconfig
     ```
   - For CentOS:
     ```
     export PKG_CONFIG_PATH=$PKG_CONFIG_PATH:/opt/mellanox/dpdk/lib64/pkgconfig
     ```
4. In order to run the application, the RegEx compiled rule files must be supplied to the app. These files usually end with `*.rof2.binary`. To compile the example rules file, run:
   ```
   cd /opt/mellanox/doca/examples/file_scan/bin/
   rxpc -f regex_rules.txt -p 0.01 -o /tmp/regex_rules
   ```
   The results of the `rxpc` are written to `/tmp/` directory, each file with the prefix `regex_rules`.

   Note: For more information, refer to NVIDIA RXP Compiler.
5. Pre-run setup.

   a). The file scan example is based on DPDK libraries. Therefore, the user is required to provide DPDK flags, and allocate huge pages. Run:
   ```
   echo 2048 > /sys/kernel/mm/hugepages/hugepages-2048kB/nr_hugepages
   ```
   b). Make sure the regex engine is active:
   ```
   systemctl status mlx-regex
   ```

If the status is inactive (`Active: failed`), run:

```
systemctl start mlx-regex
```

6. To run the application:

```
Usage: doca_file_scan [DPDK Flags] -- [DOCA Flags] [Program Flags]

DOCA Flags:
  -h, --help                Print a help synopsis
  -l, --log-level           Set the log level for the app <CRITICAL=0, DEBUG=4>

Program Flags:
  -r, --rules <path>        Path to precompiled rules file (rof2.binary)
  -d, --data <path>         Path to data file
```

Note: The RegEx accelerator explicitly requires to be initialized on p0.

For example, assuming PCIe address of the DPU is 03:00:

```
cd /opt/mellanox/doca/examples/file_scan/bin/
./doca_file_scan -a 03:00.0,class=regex -- --rules /tmp/regex_rules.rof2.binary
 --data data_to_scan.txt
```

Using a JSON file:

```
doca_file_scan --json [json_file]
```

For example:

```
/opt/mellanox/doca/examples/file_scan/bin/doca_file_scan --json /root/
file_scan_params.json
```

For additional information on available flags for DPDK, use `-h` before the `--` separator:

```
/opt/mellanox/doca/examples/file_scan/bin/doca_file_scan -h
```

For additional information on the app, use `-h` after the `--` separator:

```
/opt/mellanox/doca/examples/file_scan/bin/doca_file_scan -- -h
```

The file scan app prints the results to the standard output.

Chapter 6. Arg Parser DOCA Flags

For more information, refer to NVIDIA DOCA Arg Parser User Guide.

Flag Type	Short Flag	Long Flag/JSON Key	Description	JSON Content
DPDK Flags	a	devices	Add a PCIe device into the list of devices to probe.	`"devices": [ {"device": "regex", "id": "03: 00.0"} ]`
General Flags	l	log-level	Sets the log level for the application: ▶ CRITICAL=0 ▶ ERROR=1 ▶ WARNING=2 ▶ INFO=3 ▶ DEBUG=4	`"log-level": 4`
	h	help	Print a help synopsis	N/A
Program Flags	r	rules	Path to precompiled rules file (rof2.binary)	`"rules": "/tmp/ regex_rules.rof2.binary"`
	d	data	Path to data file	`"data": "/tmp/ data_to_scan.txt"`

Chapter 7. Running Application on Host

Host execution examples:

```
cd /opt/mellanox/doca/examples/file_scan/bin/
./doca_file_scan -a 21:00.0,class=regex -- --rules /tmp/regex_rules.rof2.binary --
data data_to_scan.txt
```

Refer to section "Running DOCA Application on Host" in NVIDIA DOCA Virtual Functions User Guide.

Chapter 8. References

- /opt/mellanox/doca/examples/file_scan/src/file_scan.c
- /opt/mellanox/doca/examples/file_scan/src/regex_rules.txt
- /opt/mellanox/doca/examples/file_scan/src/data_to_scan.txt

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. The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a world¬wide basis. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2022 NVIDIA Corporation & affiliates. All rights reserved.

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