NVIDIA DOCA DPA Execution Unit Management Tool

User Guide

MLNX-15-060583 _v2.2.0 | October 2023

Table of Contents

Chapter 1. Introduction 1

Chapter 2. Execution Unit Objects 3

Chapter 3. dpaeumgmt Commands 4

3.1. General Commands 4

3.2. Execution Unit Group Commands 4

3.2.1. EU Group Command Flags and Arguments 4

3.2.2. Info EU Group 5

3.2.3. Create EU Group 5

3.2.4. Destroy EU Group 6

3.2.5. Query EU Group 6

3.2.6. Apply EU Group 6

3.3. EU Partition Commands 7

3.3.1. EU Partition Command Flags and Arguments 7

3.3.2. Info EU Partition 7

3.3.3. Create EU Partition 8

3.3.4. Destroy EU Partition 8

3.3.5. Query EU Partition 8

Chapter 4. vHCAs and Partitions 9

Chapter 5. Known Limitations 10

Chapter 1. Introduction

This table introduces important terms for understanding this document:

Term	Definition
DPA	Data-path accelerator; an auxiliary processor designed to accelerate data-path operations.
DPA partition manager	PCIe device function capable of controlling the entire system's EUs. On NVIDIA® BlueField®-3 it is the ECPF. The DPA partition manager is by default associated with the default partition.
EU	Hardware execution unit; a logical DPA processing unit.
EU group	Collection/subset of EUs which could be created using `dpaeumgmt`. EU groups are created under an EU partition and could only be formed from the pool of EUs under that partition.
EU object	EU partition or EU group.
EU partition	An isolated pool of EUs which may be created using `dpaeumgmt`. Only when a partition is created and associated with other vHCAs are they able to use hardware resources and execute a DPA software thread.
EU affinity	The method by which a DPA thread is paired with a DPA EU. DPA supports three types of affinity: ▶ `none` – selects an EU from a pool of all available EUs ▶ `strict` – select only the specified EU (by ID) ▶ `group` – select an EU from all the EUs in the specified group

The DPA EU management tool can run either on the host machine or on the target DPU and allows users to manage the DPA's EUs which are the basic resource of the DPA. The

tool enables the resource control of EUs to optimize computation resources usage of the DPA before using DOCA FlexIO SDK API.

Without EU allocation, a DPA software thread would lack access to the hardware pipeline/CPU time resource, and consequently not be able to execute.

`dpaeumgmt` serves the following main usages:

- Running a DPA software thread with `strict` affinity on a DPA EU (i.e., running a DPA thread using only the specific preselected EU). For this purpose, `dpaeumgmt` provides an option to query the maximum EU ID allowed to use.
- Allowing a DPA software thread to run over a DPA EU from a group of EUs:
  - Once a EU group is created, it is allocated a subset of EUs.
  - `dpaeumgmt` provides an ID to the created group which can be used to run DPA applications with `group` affinity where the affinity ID would be the same as that group's ID.
- EU partition management - the ability to manage EU partitions.

When the software stack wishes to run a DPA thread with `group` affinity type, one of the available EUs from the group's collection is used for the execution.

Note: A DPA thread may execute if and only if there is an available EU for it.

Chapter 2. Execution Unit Objects

Upon boot, a default EU partition is automatically created. The default EU partition possesses all the system's EUs. The DPA partition manager function is the only function that belongs to it and can therefore control the entire resources of the system.

When running a DPA thread with `none` affinity, the EU chosen for the DPA thread to run with comes from the partition's pool of EUs. Namely, from the EUs belonging only to the DPA device's current partition which were not assigned to any EU groups (on the current partition). If the aforementioned group of EUs (i.e., the partition's default EU group) is empty, the DPA thread would fail to run with `none` affinity.

Chapter 3. dpaeumgmt Commands

`dpaeumgmt` enables users to create, destroy, and query EU objects.

Note: `dpaeumgmt` tool must run with root privileges and users must execute `sudo mst start` before using it.

Top-level `dpaeumgmt` command syntax:

```
Usage: dpaeumgmt {help|version|eu_group|partition}

Type "./dpaeumgmt help" for detailed help
```

3.1. General Commands

- Print basic usage information for the tool:
  ```
  dpaeumgmt -h
  ```
- Print a detailed help menu of the tool's commands:
  ```
  dpaeumgmt help
  ```
- Print version information:
  ```
  dpaeumgmt version
  ```

3.2. Execution Unit Group Commands

The commands listed in the following subsections are used to configure EU groups.

3.2.1. EU Group Command Flags and Arguments

The following table lists the flags relevant for `eu_group` commands. Arguments for the flags must be used within quotes (if more than one) and without extra spaces.

Short Option	Long Option	Description
`-h`	`--help`	Print out basic tool usage information.
`-d`	`--dpa_device`	The device interface name (MST/PCI/RDMA/NET).
`-r`	`--range_eus`	The range of EUs to allocate an EU group or a partition. The

Short Option	Long Option	Description
		argument must be provided within quotes.
-g	--id_group	Group ID number. This number must be positive and less than or equal to the max_num_dpa_eu_group parameter which may be retrieved using the command eu_group info -d <device>.
-n	--name_group	Group name; 15-character string. The argument must be provided within quotes.
-f	--file_groups	Full path or only the filename if it is located in the same directory as the executable directory (where dpaeumgmt is).

3.2.2. Info EU Group

Print information on the relevant DPA resources for the EU groups:

```
dpaeumgmt eu_group info --dpa_device <device>
```

Example:

```
$ sudo ./dpaeumgmt eu_group info -d mlx5_0
Max number of DPA EU groups: 15
Max number of DPA EUs in one DPA EU group: 190
Max DPA EU number available to use: 190
Max EU group name length is 15 chars
```

3.2.3. Create EU Group

Create an EU group with the specified name on the provided device's partition. The EUs indicated by the range are taken from the DPA device's EU partition.

```
dpaeumgmt eu_group create --dpa_device <device> --name_group <name> --range_eus
 <range>
```

Example:

```
$ sudo ./dpaeumgmt eu_group create -d mlx5_0 -n "HG hello world1" -r "6-8,16,55,70"
Group created successfully-
EU group ID: 1
EU group name: HG hello world
Member EUs are: 6,7,8,16,55,70
```

Note: After successfully creating an EU group, users can run a DPA thread using group affinity with the affinity type set to the group's ID.

3.2.4. Destroy EU Group

Destroy a EU group that exists on the device's partition with either the provided group name or ID.

```
dpaeumgmt eu_group destroy --dpa_device <device> [--name_group <name> | --id_group
 <id>]
```

Example:

```
$ sudo ./dpaeumgmt eu_group destroy -d mlx5_0 -g 1
Group with group id: 1, was destroyed successfully
```

3.2.5. Query EU Group

Query EU groups residing on the provided device's partition. If one of the optional parameters is used, the command only queries the specific group and prints it if it exists:

```
dpaeumgmt eu_group query --dpa_device <device> [--name_group <name> | --id_group
 <id>]
```

Example:

```
$ sudo ./dpaeumgmt eu_group query -d mlx5_0
1) EU group ID: 1
EU group name: HG hello world
Member EUs are: 6,7,8,16,55,70

In total there are 1 EU groups configured.
```

More options:

```
$ sudo ./dpaeumgmt eu_group query -d mlx5_0 -n "HG hello world"
$ sudo ./dpaeumgmt eu_group query -d mlx5_0 -g 1
```

3.2.6. Apply EU Group

Apply the EU groups provided in the file on the device's partition:

```
dpaeumgmt eu_group apply --dpa_device <device> --file_groups <file>
```

File format example:

```
{
        "eu_groups": [
                { "name": "hg1", "range": "178-180"},
                { "name": "hg2", "range": "2-10"}
        ]
}
```

Note: The command removes all the previous EU groups defined on the EU partition that the DPA device belongs to and applies the ones from the file.

Example:

```
$ sudo ./dpaeumgmt eu_group apply -d mlx5_0 --file_groups example.json
1) EU group ID: 1
EU group name: hg1
Member EUs are: 178,179,180

1) EU group ID: 2
EU group name: hg2
Member EUs are: 2,3,4,5,6,7,8,9,10

In total there are 2 EU groups configured.
```

3.3. EU Partition Commands

The commands listed in the following subsections are used to configure EU partitions.

3.3.1. EU Partition Command Flags and Arguments

The following table lists the flags relevant for EU `partition` commands. Arguments for the flags must be used within quotes (if more than one) and without extra spaces.

Short Option	Long Option	Description
`-h`	`--help`	Print out basic tool usage information.
`-d`	`--dpa_device`	The device interface name (MST/PCI/RDMA/NET).
`-r`	`--range_eus`	The range of EUs to allocate an EU group or a partition. The argument must be provided within quotes.
`-p`	`--id_partition`	Partition ID number. This number must be positive and less than or equal to the value of `max_num_dpa_eu_partition` which may be retrieved using the command `partition info -d <device>`.
`-v`	`--vhca_list`	The vHCA IDs to be associated with the partition. The argument must be provided within quotes.
`-m`	`--max_num_eu_group`	The number of EU groups to reserve for the partition upon its creation.

3.3.2. Info EU Partition

Print the relevant DPA resources of the EU partitions:

```
dpaeumgmt partition info --dpa_device <device>
```

Example:

```
$ sudo ./dpaeumgmt partition info -d mlx5_0
Max number of DPA EU partitions: 15
Max number of VHCAs associated with a single partition: 32
Max number of DPA EU groups: 15
Note- an allocation of a partition consumes from the number of DPA EU *groups*
 available to create
Max DPA EU number available to use: 190
```

3.3.3. Create EU Partition

Create an EU partition on the DPA device:

```
dpaeumgmt partition create --dpa_device <device> --vhca_list <id_list> --range_eus
 <range> --max_num_eu_group <max_num>
```

Example:

```
$ sudo ./dpaeumgmt partition create -d mlx5_0 -v 1 -r 10-20 -m 2
Partition created successfully-
EU Partition ID: 1
Maximal number of groups: 2
The partition has a total of 1 associated VHCA IDs, namely: 1
Partition's member EUs are: 10,11,12,13,14,15,16,17,18,19,20
```

3.3.4. Destroy EU Partition

Destroy an EU partition that exists on the device's partition:

```
dpaeumgmt partition destroy --dpa_device <device> --id_partition <id>
```

Example:

```
$ sudo ./dpaeumgmt partition destroy -d mlx5_0 -p 1
Partition with partition id: 1, was destroyed successfully
```

3.3.5. Query EU Partition

Query EU partitions that reside on the provided device's partition and print out the partition if it exists:

```
dpaeumgmt partition query --dpa_device <device> [--id_partition <id>]
```

Example:

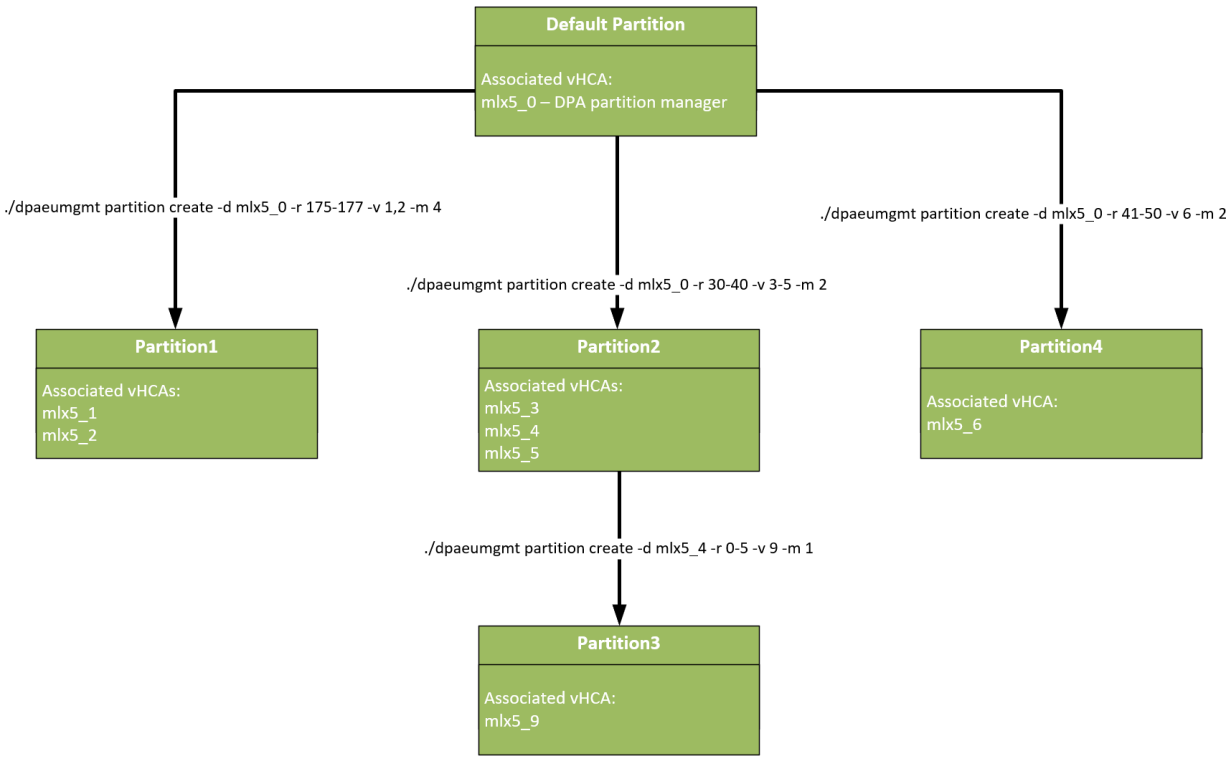
```
$ sudo ./dpaeumgmt partition query -d mlx5_0 -p 1
EU Partition ID: 1
Maximal number of groups: 2
The partition has a total of 1 associated VHCA IDs, namely: 1
Partition's member EUs are: 10,11,12,13,14,15,16,17,18,19,20
```

More options:

```
$ sudo ./dpaeumgmt partition query -d mlx5_0
```

Chapter 4. vHCAs and Partitions

The following diagram illustrates the ownership and control of a partition by a vHCA and also which vHCAs have claim to (i.e., can use) a partition.



Chapter 5. Known Limitations

- Currently, `dpaeumgmt` is only supported on the DPU not the host
- `dpaeumgmt` should run before creating a DPA process so all resources are configured ahead of time
  - Running the tool over a device with an existing DPA process results in failure
- The EU group name assigned by the user must be unique for every EU group on a specific partition or the EU group create command fails
- The creation of an EU partition consumes from the number of EU groups allowed on the vHCA's partition it is created on:
  - 1 group for the partition itself due to a default group created for each partition
  - `<max_num>` of groups which is the user's input provided upon partition creation
- Creating groups or running DPA threads in general (with any affinity) on interfaces other than ECPF, requires a configuration of a valid partition for the specific vHCA
- Only the default partition is exposed to the real EU numbers, all other partitions the user creates use virtual EUs
  - For example, if a user creates a partition with the range of EUs 20-40, querying the partition info from one of its virtual HCAs (vHCAs) would display EUs from 0-20. Therefore, the EU whose real number is 39 in this example would correspond to the virtual EU number 19.
- Group IDs on a non-default partition are virtual.
  - Different partitions can have completely distinct groups, even if they have the same ID.
  - The affinity ID parameter, specified on the FlexIO API, can distinguish between the groups according to the vHCA and application it is running on.
- vHCA ID overlap is not allowed on EU partitions
- It is not possible to query vHCA IDs with `dpaeumgmt`, these are assumed to be known by the user beforehand
- Partition destruction fails if there are EU objects that exist on that partition
- It is not possible to know which EU has been chosen to run on
- Every vHCA sees the partition it belongs to, and its resources, as the entire world. It only sees:

  - Groups and partitions it created
  - The number of EUs it was given
  - The `max_num_eu_group` of the partition it belongs to
- No guarantee regarding EU group ID that will be given on group creation
- The default groups (of every partition) cannot be managed by the user
- The EU numbers available are between 0 and the max DPA EU number available to use minus 1 (the upper limit can be queried using the info command specified above)
- `dpaeumgmt` does not support virtual functions (VFs)
- It is not possible to create partitions on other vHCAs other than the DPA partition manager function
- There are at most 16 hardware EU group entities

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

© 2023 NVIDIA Corporation & affiliates. All rights reserved.

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