

NAME

devlink-rate – devlink rate management

SYNOPSIS

devlink [*OPTIONS*] **port function rate** { *COMMAND* | **help** }

OPTIONS := { -j[*son*] | -p[*retty*] | -i[*ec*] }

devlink port function rate show [{ *DEV/PORT_INDEX* | *DEV/NODE_NAME* }]

devlink port function rate set { *DEV/PORT_INDEX* | *DEV/NODE_NAME* } [**tx_share** *VALUE*] [**tx_max** *VALUE*] [**tx_priority** *N*] [**tx_weight** *N*] [**tc-bw** *INDEX:N*] [{ **parent** *NODE_NAME* | **noparent** }]

devlink port function rate add *DEV/NODE_NAME* [**tx_share** *VALUE*] [**tx_max** *VALUE*] [**tx_priority** *N*] [**tx_weight** *N*] [**tc-bw** *INDEX:N*] [{ **parent** *NODE_NAME* | **noparent** }]

devlink port function rate del *DEV/NODE_NAME*

devlink port function rate help

DESCRIPTION**devlink port function rate show - display rate objects.**

Displays specified rate object or, if not specified, all rate objects. Rate object can be presented by one of the two types:

- leaf** Represents a single devlink port; created/destroyed by the driver and bound to the devlink port. As an example, some drivers may create a leaf rate object for every devlink port associated with VF. Since leaves have a 1-to-1 mapping to their devlink port, in user space they are referred to as corresponding devlink port *DEV/PORT_INDEX*;
- node** Represents a group of rate objects; created/deleted by the user (see command below) and bound to the devlink device rather than to the devlink port. In userspace it is referred as *DEV/NODE_NAME*, where node name can be any, except decimal number, to avoid collisions with leaves.

Command output show rate object identifier, its type and rate values along with parent node name. Rate values printed in SI units which are more suitable to represent specific value. To print values in IEC units -i switch is used. JSON (-j) output always print rate values in bytes per second. Zero rate values means "unlimited" rates and omitted in output, as well as parent node name.

devlink port function rate set - set rate object parameters.

Allows set rate object's parameters. If any parameter specified multiple times the last occurrence is used.

DEV/PORT_INDEX - specifies devlink leaf rate object.

DEV/NODE_NAME - specifies devlink node rate object.

tx_share *VALUE* - specifies minimal tx rate value shared among all rate objects. If rate object is a part of some rate group, then this value shared with rate objects of this rate group.

tx_max *VALUE* - specifies maximum tx rate value.

tx_priority *N* - allows for usage of strict priority arbiter among siblings. This arbitration scheme attempts to schedule nodes based on their priority as long as the nodes remain within their bandwidth limit. The higher the priority the higher the probability that the node will get selected for scheduling.

tx_weight *N* - allows for usage of Weighted Fair Queuing arbitration scheme among siblings. This arbitration scheme can be used simultaneously with the strict priority. As a node is configured with a higher rate

it gets more BW relative to it's siblings. Values are relative like a percentage points, they basically tell how much BW should node take relative to its siblings.

tc-bw INDEX:N - allows the user to assign relative bandwidth shares to specific traffic classes using the IEEE 802.1Qaz standard. The values determine how bandwidth is distributed between traffic classes in proportion to one another. If not specified, the default bandwidth allocation is applied.

VALUE These parameter accept a floating point number, possibly followed by either a unit (both SI and IEC units supported).

bit or a bare number

Bits per second

kbit Kilobits per second

mbit Megabits per second

gbit Gigabits per second

tbit Terabits per second

bps Bytes per second

kbps Kilobytes per second

mbps Megabytes per second

gbps Gigabytes per second

tbps Terabytes per second

To specify in IEC units, replace the SI prefix (k-, m-, g-, t-) with IEC prefix (ki-, mi-, gi- and ti-) respectively. Input is case-insensitive.

INDEX These parameters represent the traffic class index in the *tc-bw* option. The traffic class is specified as an integer value, ranging from 0 to 7, which maps to the defined traffic classes under the IEEE 802.1Qaz standard.

N These parameters accept an integer meaning weight or priority of a node.

parent NODE_NAME | noparent - set rate object parent to existing node with name *NODE_NAME* or unset parent. Rate limits of the parent node are applied to all its children. Actual behavior is a detail of the driver's implementation. Setting parent to empty ("") name due to the kernel logic is treated as unsetting the parent.

devlink port function rate add - create node rate object with specified parameters.

Creates rate object of type node and sets parameters. Parameters same as for the "set" command.

DEV/NODE_NAME - specifies the devlink node rate object to create.

devlink port function rate del - delete node rate object

Delete specified devlink node rate object. Node can't be deleted if there is any child, user must explicitly unset the parent.

DEV/NODE_NAME - specifies devlink node rate object to delete.

devlink port function rate help - display usage information

Display devlink rate usage information

EXAMPLES

* Display all rate objects:

```
# devlink port function rate show
pci/0000:03:00.0/1 type leaf parent some_group
```

```
pci/0000:03:00.0/2 type leaf tx_share 12Mbit
pci/0000:03:00.0/some_group type node tx_share 1Gbps tx_max 5Gbps
```

- * Display leaf rate object bound to the 1st devlink port of the pci/0000:03:00.0 device:

```
# devlink port function rate show pci/0000:03:00.0/1
pci/0000:03:00.0/1 type leaf
```

- * Display leaf rate object rate values using IEC units:

```
# devlink -i port function rate show pci/0000:03:00.0/2
pci/0000:03:00.0/2 type leaf 11718Kibit
```

- * Display node rate object with name some_group of the pci/0000:03:00.0 device:

```
# devlink port function rate show pci/0000:03:00.0/some_group
pci/0000:03:00.0/some_group type node
```

- * Display pci/0000:03:00.0/2 leaf rate object as pretty JSON output:

```
# devlink -jp port function rate show pci/0000:03:00.0/2
{
  "rate": {
    "pci/0000:03:00.0/2": {
      "type": "leaf",
      "tx_share": 1500000
    }
  }
}
```

- * Create node rate object with name "1st_group" on pci/0000:03:00.0 device:

```
# devlink port function rate add pci/0000:03:00.0/1st_group
```

- * Create node rate object with specified parameters:

```
# devlink port function rate add pci/0000:03:00.0/2nd_group \
tx_share 10Mbit tx_max 30Mbit parent 1st_group
```

- * Set parameters to the specified leaf rate object:

```
# devlink port function rate set pci/0000:03:00.0/1 \
tx_share 2Mbit tx_max 10Mbit
```

- * Set leaf's parent to "1st_group":

```
# devlink port function rate set pci/0000:03:00.0/1 parent 1st_group
```

- * Unset leaf's parent:

```
# devlink port function rate set pci/0000:03:00.0/1 noparent
```

- * Delete node rate object:

```
# devlink port function rate del pci/0000:03:00.0/2nd_group
```

DEVLINK-RATE(8)

Linux

DEVLINK-RATE(8)

SEE ALSO

devlink(8), devlink-port(8)

AUTHOR

Dmytro Linkin <dlinkin@nvidia.com>