Nokia Validated Design

AMD Pensando Pollara Validation

3HE-21917-AAAA-TQZZA
Issue 1
February 2026

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

Nokia is committed to diversity and inclusion. We are continuously reviewing our customer documentation and consulting with standards bodies to ensure that terminology is inclusive and aligned with the industry. Our future customer documentation will be updated accordingly.

This document includes Nokia proprietary and confidential information, which may not be distributed or disclosed to any third parties without the prior written consent of Nokia.

This document is intended for use by Nokia's customers ("You"/"Your") in connection with a product purchased or licensed from any company within Nokia Group of Companies. You agree to notify Nokia of any errors you may find in this document; however, should you elect to use this document for any purpose(s) for which it is not intended, You understand and warrant that any determinations You may make or actions You may take will be based upon Your independent judgment and analysis of the content of this document.

Nokia reserves the right to make changes to this document without notice.

No part of this document may be copied, reproduced, modified, or transmitted.

NO WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY OF AVAILABILITY, ACCURACY, RELIABILITY, TITLE, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, IS MADE IN RELATION TO THE CONTENT OF THIS DOCUMENT. IN NO EVENT WILL NOKIA BE LIABLE FOR ANY DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, DIRECT, INDIRECT, INCIDENTAL OR CONSEQUENTIAL OR ANY LOSSES, SUCH AS BUT NOT LIMITED TO LOSS OF PROFIT, REVENUE, BUSINESS INTERRUPTION, BUSINESS OPPORTUNITY OR DATA THAT MAY ARISE FROM THE USE OF THIS DOCUMENT OR THE INFORMATION IN IT, EVEN IN THE CASE OF ERRORS IN OR OMISSIONS FROM THIS DOCUMENT OR ITS CONTENT.

Copyright and trademark: Nokia is a registered trademark of Nokia Corporation. Other product names mentioned in this document may be trademarks of their respective owners.

© 2026 Nokia.

Contents

1 Executive summary

This document provides validation for the AMD Pensando Pollara network interface cards when used with the Nokia Validated Design (NVD) described in *AI Cluster with Backend, Frontend, and Storage Networks*, which is available on the Nokia Documentation Center. This AI Data Center NVD solution comprises AMD MI300X GPUs hosted in SMCI servers and connected to Nokia 7220-IXR-H4 and 7220-IXR-H5 switches in a rail optimized design.

The solution described in the NVD was tested based on Broadcom Thor 2 network interface cards on the AMD GPU servers and was validated end-to-end.

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

The objective of this document is to provide a network-level validation with performance tests for the same solution using AMD Pensando Pollara network interface cards.

2 Reference architecture

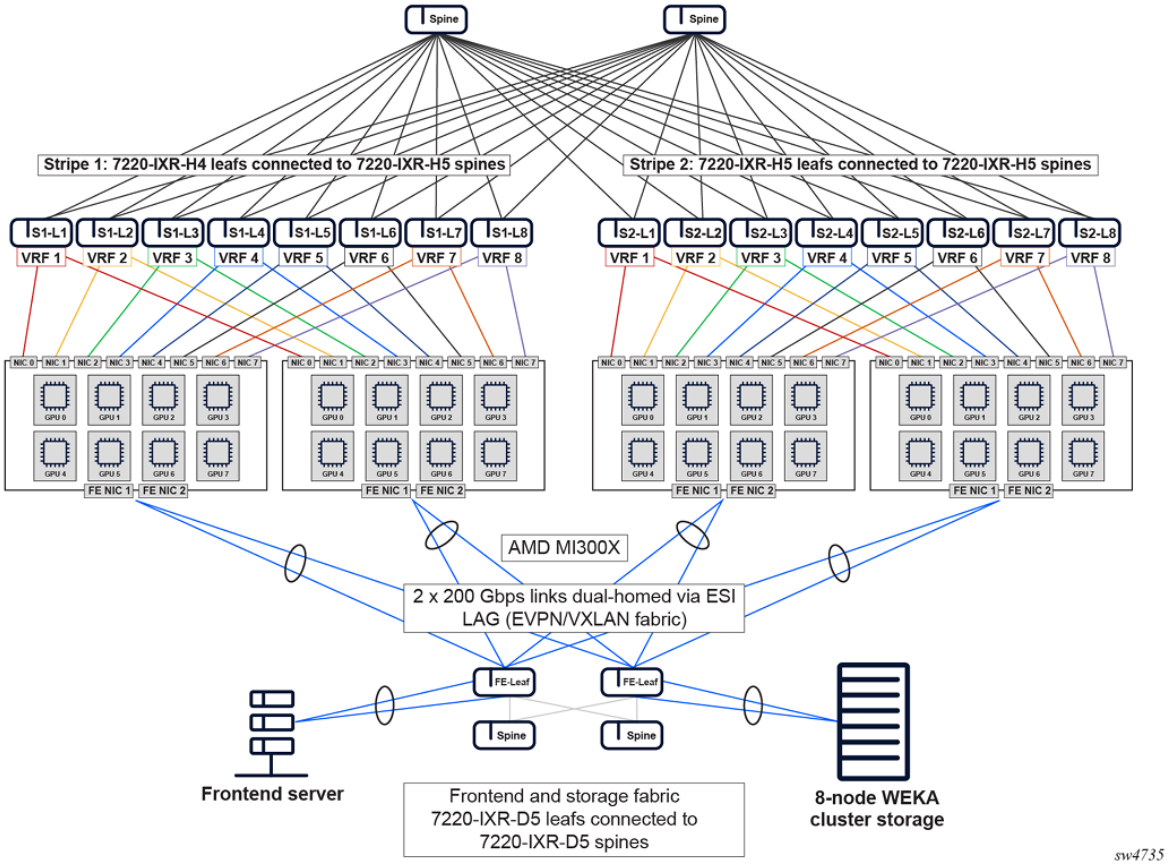

Figure 1. **HLD – Nokia hybrid training and inference cluster design**

The setup has four SMCI servers with AMD MI300X GPUs as shown above (slate 1, 2, 3, 4). For the purposes of this validation, the GPU NICs of slates 1, 2, 3 and 4 on the H5 fabric have been replaced with AMD Pensando Pollara NICs.

Scope of validation

The scope of this test and validation is limited to:

- hardware validation
- network connectivity

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

- multinode RoCEv2 performance tests
- multinode RCCL performance tests

3 Hardware and optics

The following table shows the hardware that has been used in the validation process. The selection of optics was determined by laboratory availability.

Connection					Connectivity option
Endpoints			Speed	Port type	AOC or transceiver and cable
1	A	AMD Pensando Pollara 400 AI NIC	400G	QSFP112	AOC: • Eoptolink • EOLD-8HG-PCT-05C1 • 800G QSFPDD to 2 x 400G QSFP112 Breakout AOC 5m
	B	7220 IXR-H5-32D	400G	QSFP-DD	
2	A	AMD Pensando Pollara 400 AI NIC	400G	QSFP112	AOC: • Eoptolink • EOLD-8HG-PCT-05C1 • 800G QSFPDD to 2 x 400G QSFP112 Breakout AOC 5m
	B	7220 IXR-H5-64D	400G	QSFP-DD	
3	A	AMD Pensando Pollara 400 AI NIC	400G	QSFP112	Transceiver: • 3HE19539AARA01 • QSFP112 – DR4 400G Cable: MPO-12
	B	7220 IXR-H4-32D	400G	QSFP-DD	Transceiver: • 3HE15271AARA01 • QSFP-DD – DR4 400G Cable: MPO-12

4 AMD Pensando Pollara setup

This section outlines the detailed installation procedure for the required software tools, drivers, and firmware, which are essential for managing and configuring AMD Pensando Pollara NICs. These necessary firmware and software bundles can be obtained from the AMD Pensando Customer Portal, located at pensandosupport.amd.com. These come as a package, usually in the form ainic_bundle_*version*.tar.gz.

The following software versions have been used:

Software	Version
AI NIC bundle	ainic_bundle_1.117.5-a-56.tar.gz
Host tool nicctl	1.117.5-a-56
Ethernet and RDMA drivers (ionic driver)	25.12.15.001
ROCm	6.4.3

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

RCCL	2.22.3
MPI	1.15.0

4.1 Installation

Prerequisites:

- Root access is required to complete the installation.
- Write access by root is required in the location where the tarball is unpacked.

Unpack the ainic_bundle_1.117.1-a-63.tar.gz package.

```
root@slate1:~# tar -xvf ainic_bundle_1.117.1-a-63.tar.gz
ainic_bundle_1.117.1-a-63/
ainic_bundle_1.117.1-a-63/host_sw_pkg.tar.gz
ainic_bundle_1.117.1-a-63/firmware/
ainic_bundle_1.117.1-a-63/firmware/ainic_fw_salina_secure.pldmfw
ainic_bundle_1.117.1-a-63/firmware/ainic_fw_salina_secure.tar
ainic_bundle_1.117.1-a-63/firmware/ainic_fw_salina.pldmfw
ainic_bundle_1.117.1-a-63/firmware/gold/
ainic_bundle_1.117.1-a-63/firmware/gold/ainic_gold_fw_salina_secure.tar
ainic_bundle_1.117.1-a-63/firmware/gold/ainic_gold_fw_salina_secure.pldmfw
ainic_bundle_1.117.1-a-63/firmware/gold/ainic_gold_fw_salina.tar
ainic_bundle_1.117.1-a-63/firmware/gold/ainic_gold_fw_salina.pldmfw
ainic_bundle_1.117.1-a-63/firmware/ainic_fw_salina.tar
ainic_bundle_1.117.1-a-63/README.md
root@slate1:~#
```

The unpacked directory contains a firmware directory, which holds the NIC software images and the host_sw_pkg.tar.gz file, providing the host tools and the ionic driver.

```
root@slate1:~# cd ainic_bundle_1.117.1-a-63
root@slate1:~/ainic_bundle_1.117.1-a-63# ll
total 511852
drwxr-xr-x 3 root root      4096 Oct  4 12:56 ./
drwx------ 6 root root      4096 Jan  6 17:11 ../
drwxr-xr-x 3 root root      4096 Oct  4 12:56 firmware/
-rw-r--r-- 1 root root 524041831 Oct  4 12:56 host_sw_pkg.tar.gz
-rw-r--r-- 1 root root     74950 Oct  4 12:56 README.md
```

4.2 Driver installation

Perform the following steps.

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

1. Extract host_sw_pkg.tar.gz, then navigate to the host_sw_pkg directory.
2. Run the install.sh script to install the ionic driver and host utilities, such as nicctl. The nicctl tool is used to manage the NIC lifecycle.

```
root@slate1:~/ainic_bundle_1.117.1-a-63/host_sw_pkg# ./install.sh
INFO: Handling compatibility for module 'nicctl'
INFO: Handling compatibility for module 'pds_driver_dkms'
INFO: Handling compatibility for module 'ipc_driver_dkms'
INFO: Handling compatibility for module 'ionic_driver_dkms'
INFO: Handling compatibility for module 'rdma_core'
INFO: Handling compatibility for module 'perftest'
…
<snipped>
…
INFO: Tool version setup completed for version 1.117.1-a-63.
WARNING: Please ensure the following environment variables are correctly set:
WARNING:     1. PATH: /usr/sbin must be included.
WARNING:     2. LD_LIBRARY_PATH: /usr/lib must be included.
WARNING: If these paths are missing, tools and libraries may not function correctly.
INFO: Installation completed successfully
root@slate1:~/ainic_bundle_1.117.1-a-63/host_sw_pkg#
```

4.3 Firmware installation

Use the nicctl tool to update the firmware. You can update a single card or all cards using the commands shown in the following steps, where `card_uuid (-c)` is the unique identifier for a NIC card in the system. If no card is specified, all cards are upgraded by default.

Perform the following steps.

1. To view the card_uuid values for all cards in the system, run the following command:

```
root@slate1:~/ainic_bundle_1.117.1-a-63# nicctl show card
-------------------------------------------------------------------------------------
Id                                      PCIe BDF       ASIC      F/W partition Serial number
-------------------------------------------------------------------------------------
42424650-4c32-3532-3230-383238000000    0000:06:00.0   salina    A             FPL25220828
42424650-4c32-3532-3330-324430000000    0000:23:00.0   salina    A             FPL252302D0
42424650-4c32-3532-3230-374434000000    0000:43:00.0   salina    A             FPL252207D4
42424650-4c32-3532-3230-413439000000    0000:66:00.0   salina    A             FPL25220A49
42424650-4c32-3532-3230-344433000000    0000:86:00.0   salina    A             FPL252204D3
42424650-4c32-3532-3230-363738000000    0000:a3:00.0   salina    A             FPL25220678
42424650-4c32-3532-3330-333942000000    0000:c3:00.0   salina    A             FPL2523039B
42424650-4c32-3532-3330-343542000000    0000:e6:00.0   salina    A             FPL2523045B
```

2. Use nicctl to update the firmware by pointing to the ainic_fw_salina.tar package in the ainic_bundle_1.117.1-a-63/firmware directory. The following example shows how to upgrade a single NIC.

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
root@slate1:~/ainic_bundle_1.117.1-a-63/firmware# nicctl update firmware -i ainic_fw_salina.tar -c
42424650-4c32-3532-3230-383238000000
--------------------------------------------------------------------------------
Card Id                                 Stage                                   Progress
--------------------------------------------------------------------------------
42424650-4c32-3532-3230-383238000000    Done                                    100% [02:59.412]
NIC 42424650-4c32-3532-3230-383238000000 (0000:06:00.0) : Successful
```

3. After the cards are upgraded, reset them to complete upgrade and boot to the new version. You can reset one card or all cards.

```
root@slate1:~/ainic_bundle_1.117.1-a-63/firmware# nicctl reset card -c 42424650-4c32-3532-3230-
383238000000
NIC 42424650-4c32-3532-3230-383238000000 (0000:06:00.0) : Card reset triggered, wait for
completion (64 secs)
NIC 42424650-4c32-3532-3230-383238000000 (0000:06:00.0) : Card reset successful
root@slate2:~/ainic_bundle_1.117.1-a-63/firmware#
```

4.4 Verify software installation

Perform the following steps.

1. Verify successful installation of the host software tools and RDMA driver.

```
cse@slate1:~$ sudo nicctl show version host-software
[sudo] password for cse:
nicctl                          : 1.117.5-a-56
IPC driver                      : 1.117.5.a.56
ionic driver                    : 25.12.15.001
```

2. Verify successful firmware installation.

```
cse@slate1:~$ sudo nicctl show version firmware | grep -i firmware
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
Firmware-A                      : 1.117.5-a-56
```

3. Verify that the RDMA interfaces are operational UP.

```
cse@slate1:~$ rdma link show
link ionic_0/1 state ACTIVE physical_state LINK_UP netdev enp9s0
link ionic_1/1 state ACTIVE physical_state LINK_UP netdev enp38s0
link ionic_2/1 state ACTIVE physical_state LINK_UP netdev enp70s0
link ionic_3/1 state ACTIVE physical_state LINK_UP netdev enp105s0
link ionic_4/1 state ACTIVE physical_state LINK_UP netdev enp137s0
link ionic_5/1 state ACTIVE physical_state LINK_UP netdev enp166s0
link ionic_6/1 state ACTIVE physical_state LINK_UP netdev enp198s0
link ionic_7/1 state ACTIVE physical_state LINK_UP netdev enp233s0
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

5 Congestion management

This section explains how to configure PFC, ECN, and traffic classification on AMD Pensando Pollara NIC cards using the nicctl tool. These settings must match the configuration of the backend network. By default, the NIC cards classify all traffic as priority 0, so they must be explicitly configured. The values shown in the following table are applied.

Type of traffic	Classification DSCP	Queue	PFC	Scheduling
All other	0-25, 27-47, 49-63	0	—	DWRR: 10%
RoCEv2	26	3	Enabled	DWRR: 90%
Control traffic (CNP)	48	6	—	SPQ, rate-limit 10 Gb/s

5.1 QoS configuration

The following script ./pollara-nic-qos.sh is used to configure the classification, scheduling, and PFC.

```
#!/bin/bash

for i in $(sudo nicctl show port | grep Port | awk {'print $3'}); do sudo nicctl update port -p $i --pause-
type pfc --rx-pause enable --tx-pause enable; done

for i in $(sudo nicctl show port | grep Port | awk '{print $3}'); do sudo nicctl update port --port $i --auto-
neg enable; done

cts_dscp=48             # control traffic (CTS,CNP) DSCP value --> forwarding class 6

cts_prio=6              # control traffic priority = forwarding class 6

data_dscp=26            # RoCEv2 data DSCP value --> forwarding class 3

data_prio=3             # RoCEv2 traffic priority = forwarding class 3

default_prio=0          # all other --> forwarding class 0

cnp_dscp=48             # controll traffic (CTS,CNP) DSCP value --> forwarding class 6

cnp_prio=6              # controll traffic priority = forwarding class 6

sudo nicctl update qos pfc --priority 0 --no-drop disable

sudo nicctl update qos --classification-type dscp

sudo nicctl update qos dscp-to-priority --dscp 0-63 --priority 0

sudo nicctl update qos dscp-to-priority --dscp 0-25,27-47,49-63 --priority $default_prio

sudo nicctl update qos dscp-to-priority --dscp $cts_dscp --priority $cts_prio

sudo nicctl update qos dscp-to-priority --dscp $data_dscp --priority $data_prio
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
sudo nicctl update qos dscp-to-priority --dscp $cnp_dscp --priority $cnp_prio

sudo nicctl update qos dscp-to-purpose --dscp $cts_dscp --purpose xccl-cts

sudo nicctl update qos dscp-to-purpose --dscp $data_dscp --purpose data

sudo nicctl update qos pfc --priority $data_prio --no-drop enable

sudo nicctl update qos scheduling --priority $data_prio,$default_prio,$cts_prio --dwrr 90,10,0 --rate-limit
0,0,10
```

The pollara-qos-dcqn.sh script is used to configure DCQCN settings, including ECN CNP packet values and alpha values.

```
#!/bin/bash

TOKEN_BUCKET_SIZE=800000
AI_RATE=160
ALPHA_UPDATE_INTERVAL=1
ALPHA_UPDATE_G=512
INITIAL_ALPHA_VALUE=64
RATE_INCREASE_BYTE_COUNT=431068
HAI_RATE=300
RATE_REDUCE_MONITOR_PERIOD=1
RATE_INCREASE_THRESHOLD=1
RATE_INCREASE_INTERVAL=1
CNP_DSCP=48

ROCE_DEVICES=$(ibv_devices | grep ionic_ | awk '{print $1}' | paste -sd " ")
for roce_dev in $ROCE_DEVICES
do
    sudo nicctl update dcqcn -r "$roce_dev" -i 1 \
        --token-bucket-size "$TOKEN_BUCKET_SIZE" \
        --ai-rate "$AI_RATE" \
        --alpha-update-interval "$ALPHA_UPDATE_INTERVAL" \
        --alpha-update-g "$ALPHA_UPDATE_G" \
        --initial-alpha-value "$INITIAL_ALPHA_VALUE" \
        --rate-increase-byte-count "$RATE_INCREASE_BYTE_COUNT" \
        --hai-rate "$HAI_RATE" \
        --rate-reduce-monitor-period "$RATE_REDUCE_MONITOR_PERIOD" \
        --rate-increase-threshold "$RATE_INCREASE_THRESHOLD" \
        --rate-increase-interval "$RATE_INCREASE_INTERVAL" \
        --cnp-dscp "$CNP_DSCP"
done
```

Note: Any QoS configuration applied using nicctl is not persistent and must be reapplied after each reboot.

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

5.2 Verify the QoS configuration

Verify that the desired QoS configuration has been applied on all NIC cards according to the preceding table.

```
cse@slate1:$ sudo nicctl show qos --card 42424650-4c32-3532-3430-393535000000

NIC  : 42424650-4c32-3532-3430-393535000000 (0000:06:00.0)

Port : 0490813a-08f0-4242-4242-000011010000

  Classification type         : DSCP

  DSCP-to-priority :
    DSCP bitmap               : 0xfffeffffffbfffff ==> priority : 0
    DSCP bitmap               : 0x0000000004000000 ==> priority : 3
    DSCP bitmap               : 0x0001000000000000 ==> priority : 6
    DSCP                      : 0-25, 27-47, 49-63 ==> priority : 0
    DSCP                      : 26 ==> priority : 3
    DSCP                      : 48 ==> priority : 6

  DSCP-to-purpose             : 26 ==> data
                              : 48 ==> xccl-cts

  PFC :
    PFC priority bitmap       : 0x8
    PFC no-drop priorities    : 3

  Scheduling :
    ------------------------------------------
    Priority  Scheduling  Bandwidth Rate-limit
              Type        (in %age) (in Gbps)
    ------------------------------------------
    0         DWRR        10        N/A
    3         DWRR        90        N/A
    6         strict      N/A       10
```

If the configuration was not applied successfully, or was not applied at all, the default QoS configuration is displayed, as shown below.

```
root@slate1: # nicctl show qos --card 42424650-4c32-3532-3430-393535000000
NIC  : 42424650-4c32-3532-3430-393535000000 (0000:06:00.0)
Port : 0490813a-08f0-4242-4242-000011010000
  Classification type         : DSCP
  DSCP-to-priority :
    DSCP bitmap               : 0xffffffffffffffff ==> priority : 0
    DSCP                      : 0-63 ==> priority : 0
  PFC :
    PFC priority bitmap       : 0x0
    PFC no-drop priorities    :
  Scheduling :
    ------------------------------------------
    Priority  Scheduling  Bandwidth Rate-limit
              Type        (in %age) (in Gbps)
    ------------------------------------------
    0         DWRR        0         N/A
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

6 Performance tests

RoCEv2 perf test	
ib_write_bw average for different data paths	CPU-to-CPU – 779.68 Gb/s bidir GPU-to-GPU with peer-mem – 779.69 Gb/s bidir GPU-to-GPU with DMAbuf – 779.68 Gb/s bidir
ib_write_bw for different packet sizes 256 to 4096 bytes	256 – 73,69 Gb/s 1024 – 336.49 Gb/s 2048 – 380.35 Gb/s 4096 – 389.84 Gb/s

The following performance tests use benchmark tools such as perftest to measure NIC throughput and RDMA read/write performance between two GPU endpoints. The perftest package includes a set of microbenchmark tools, each designed to test a specific aspect of performance. These tests are typically run in a client-server setup where one GPU acts as the client and the other GPU acts as the server.

Test name	Metric measured	RDMA operation type
ib_send_bw	Bandwidth	Send
ib_send_lat	Latency	Send
ib_write_bw	Bandwidth	RDMA write
ib_write_lat	Latency	RDMA write
ib_read_bw	Bandwidth	RDMA read
ib_read_lat	Latency	RDMA read
ib_atomic_bw	Bandwidth	RDMA atomic
ib_atomic_lat	Latency	RDMA atomic

Running perftest with GPU memory requires significantly more setup than using Open MPI with RCCL, which automatically discovers the GPU and NIC topology. To use perftest, GPU and NIC PCI addresses must be manually mapped to identify which devices share the same PCI bridge. This is done by listing the RoCE interface names (ionic_*n*) associated with each NIC.

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

Refer to the *AI Cluster with Backend, Frontend, and Storage Networks NVD*, section 6.2.4.2 “GPU NIC mapping and configuration”, for detailed instructions on performing the GPU-to-NIC mapping.

NIC NUMA	NIC slot	NIC PCI address	NIC BIOS name	RoCE NIC name	Closet GPU-ID	GPU PCI address	GPU OAM-ID	Observed ROCM-ID
0	1	26:00.0	enp38s0	ionic_1	GPU1	29:00.0	6	1
0	2	09:00.0	enp9s0	ionic_0	GPU0	05:00.0	7	0
0	3	46:00.0	enp70s0	ionic_2	GPU2	49:00.0	4	2
0	4	69:00.0	enp105s0	ionic_3	GPU3	65:00.0	5	3
1	5	A6:00.0	enp166s0	ionic_4	GUP5	A9:00.0	2	5
1	6	89:00.0	enp137s0	ionic_5	GUP4	85:00.0	3	4
1	7	C6:00.0	enp198s0	ionic_6	GPU6	C9:00.0	0	6
1	8	E9:00.0	enp233s0	ionic_7	GUP7	E5:00.0	1	7

6.1 RDMA WRITE using ib_write_bw

RDMA WRITE uses the ib_write_bw tool to help measure the throughput of write operations between two NICs. The client issues an RDMA WRITE operation to memory on the server NIC.

The test can be run with memory located either in system RAM (CPU-to-CPU) or on GPU device memory (GPU-to-GPU). In CPU-to-CPU mode, both the client and server use host memory, representing the standard RDMA path. In GPU-to-GPU mode, memory resides on the GPUs and the RDMA NIC accesses it directly, bypassing the CPU. Bypassing the CPU can be done by using either peer_mem, a kernel module that exposes GPU memory to the NIC, or DMAbuf, a more portable mechanism for sharing GPU memory across devices. Both methods enable high-bandwidth, low-latency transfers directly between GPUs.

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

The following table provides an overview of all the options that have been used in conducting the performance test while using the ib_write_bw tool.

Flag	Usage
ib_write_bw	Uses RDMA write operations for measurements
-d	Sets the name of the RDMA device in use
-x	Sets the GID index to be used for the test
-s	Sets the message size in bytes
-q	Number of queue pairs to be used for the test
-D	Test duration in seconds
-b	Enables bidirectional testing
--sl	Sets the traffic class value
-F	Forces a new connection for each iteration
--reports_gbits	Reports bandwidth in Gb/s
-p	Specifies the port number to be used for connections
numactl –-cpunodebind=netdev:<interface>	Ensures the CPU cores used are from the same NUMA node as used in the interface
numactl --localloc	Ensures the memory used is from the same NUMA node
--use_rocm	Uses the selected ROCm device for GPUDirect RDMA testing, used in GPU-to-GPU mode
--use_rocm_dmabuf	Uses ROCm DMA-BUF for GPUDirect RDMA testing, used in GPU-to-GPU mode.
--ipv6	Uses IPv6 GID
--ipv6-addr	Sets the IPv6 source IP address
-m	Sets the MTU size

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

6.1.1 RDMA ib_write_bw in CPU-to-CPU mode

This test was conducted between NIC1 on server-1 (server) and NIC1 on server-2 (client), with traffic passing through the switch. The example below demonstrates how the RDMA WRITE operation was used to measure performance in CPU-to-CPU mode.

The tests were performed by first running the following command on the server side (for example, on slate1):

```
numactl --cpunodebind=netdev:enp38s0 --localalloc ./ib_write_bw -d ionic_1 -F -x 2 -q 4 --sl 3 --
report_gbits -b -p 5001 -D 60 -s 1M --ipv6 --ipv6-addr
```

Afterward, in another terminal window, we ran the following command to start the client on server-2. The specified IPv6 address is the one assigned to the interface mapped to RDMA interface ionic_1.

```
numactl --cpunodebind=netdev:enp38s0 --localalloc ./ib_write_bw -d ionic_1 -F -x 2 -q 4 -b --sl 3
-D 60 -s 1M --report_gbits -p 5001 --ipv6 --ipv6-addr fd00:2:9:1:1:1:0:2
 WARNING: BW peak won't be measured in this run.
---------------------------------------------------------------------------------------
                    RDMA_Write Bidirectional BW Test
 Dual-port       : OFF          Device         : ionic_1
 Number of qps   : 4            Transport type : IB
 Connection type : RC           Using SRQ      : OFF
 PCIe relax order: ON           Lock-free      : OFF
 ibv_wr* API     : OFF          Using DDP      : OFF
 TX depth        : 128
 CQ Moderation   : 1
 CQE Poll Batch  : Dynamic
 Mtu             : 4096[B]
 Link type       : Ethernet
 GID index       : 2
 Max inline data : 0[B]
 rdma_cm QPs     : OFF
 Use ROCm memory : OFF
 Data ex. method : Ethernet
---------------------------------------------------------------------------------------
 local address: LID 0000 QPN 0x0002 PSN 0x9a2285 RKey 0x000146 VAddr 0x007ee5de7ff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0800 PSN 0x7e0dc2 RKey 0x000146 VAddr 0x007ee5de8ff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0003 PSN 0xdb35ae RKey 0x000146 VAddr 0x007ee5de9ff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0801 PSN 0xa85b2e RKey 0x000146 VAddr 0x007ee5deaff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 remote address: LID 0000 QPN 0x0800 PSN 0x1be8dc RKey 0x000171 VAddr 0x007a87469ff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0002 PSN 0x25aacd RKey 0x000171 VAddr 0x007a8746aff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0801 PSN 0x74f55d RKey 0x000171 VAddr 0x007a8746bff000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0003 PSN 0xaa32b1 RKey 0x000171 VAddr 0x007a8746cff000
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
---------------------------------------------------------------------------------------
 #bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
 1048576    1392093          0.00               779.68               0.092945
---------------------------------------------------------------------------------------
```

6.1.2 RDMA ib_write_bw in GPU-to-GPU mode (peer_mem)

For this test, we first ran the following command to start the server on server-1 NIC1.

```
numactl --cpunodebind=netdev:enp38s0 --localalloc ./ib_write_bw -d ionic_1 -F -x 2 -q 4 --sl 3 --
report_gbits -b -p 5001 -D 60 -s 1M --use_rocm=1 --ipv6 --ipv6-addr
```

Next, we ran the command to start the client on server-2 NIC2.

```
numactl --cpunodebind=netdev:enp38s0 --localalloc ./ib_write_bw -d ionic_1 -F -x 2 -q 4 -b --sl 3
-D 60 -s 1M --report_gbits -p 5001 --use_rocm=1 --ipv6 --ipv6-addr fd00:2:9:1:1:1:0:2
 WARNING: BW peak won't be measured in this run.
Using ROCm Device with ID: 1, Name: AMD Instinct MI300X, PCI Bus ID: 0x29, GCN Arch:
gfx942:sramecc+:xnack-
allocated 8388608 bytes of GPU buffer at 0x70fa6c400000
---------------------------------------------------------------------------------------
                    RDMA_Write Bidirectional BW Test
 Dual-port       : OFF          Device         : ionic_1
 Number of qps   : 4            Transport type : IB
 Connection type : RC           Using SRQ      : OFF
 PCIe relax order: ON           Lock-free      : OFF
 ibv_wr* API     : OFF          Using DDP      : OFF
 TX depth        : 128
 CQ Moderation   : 1
 CQE Poll Batch  : Dynamic
 Mtu             : 4096[B]
 Link type       : Ethernet
 GID index       : 2
 Max inline data : 0[B]
 rdma_cm QPs     : OFF
 Use ROCm memory : ON
 Data ex. method : Ethernet
---------------------------------------------------------------------------------------
 local address: LID 0000 QPN 0x0002 PSN 0x729590 RKey 0x000145 VAddr 0x0070fa6c800000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0800 PSN 0x26eb71 RKey 0x000145 VAddr 0x0070fa6c900000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0003 PSN 0x40d331 RKey 0x000145 VAddr 0x0070fa6ca00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0801 PSN 0x341df5 RKey 0x000145 VAddr 0x0070fa6cb00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 remote address: LID 0000 QPN 0x0800 PSN 0xc21481 RKey 0x000170 VAddr 0x007bab61400000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0002 PSN 0x387e0e RKey 0x000170 VAddr 0x007bab61500000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0801 PSN 0x314e0a RKey 0x000170 VAddr 0x007bab61600000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0003 PSN 0x3625a RKey 0x000170 VAddr 0x007bab61700000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
```

```
---------------------------------------------------------------------------------------
 #bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
 1048576    1392108          0.00               779.69               0.092946
---------------------------------------------------------------------------------------
deallocating GPU buffer 0x70fa6c400000
```

6.1.3 RDMA ib_write_bw in GPU-to-GPU mode (DMABuf)

For this test, first we ran the following command to start the server on server-1 NIC1.

```
numactl --cpunodebind=netdev:enp38s0 --localalloc ./ib_write_bw -d ionic_1 -D 60 -F -x 2 -q 4 --sl
3 -s 1M --report_gbits -b -p 5001 --use_rocm=1 --use_rocm_dmabuf --ipv6 --ipv6-addr
```

Then, we ran the command to start the client on server-2 NIC2.

```
numactl --cpunodebind=netdev:enp38s0 --localalloc ./ib_write_bw -d ionic_1 -D 60 -F -x 2 -q 4 -b -
-sl 3 -s 1M --report_gbits -p 5001 --use_rocm=1 --use_rocm_dmabuf --ipv6 --ipv6-addr
fd00:2:9:1:1:1:0:2
 WARNING: BW peak won't be measured in this run.
Using ROCm Device with ID: 1, Name: AMD Instinct MI300X, PCI Bus ID: 0x29, GCN Arch:
gfx942:sramecc+:xnack-
using DMA-BUF for GPU buffer address at 0x7416c5600000 aligned at 0x7416c5600000 with aligned size
8388608
dmabuf export addr 0x7416c5600000 8388608 to dmabuf fd 8 offset 0
allocated 8388608 bytes of GPU buffer at 0x7416c5600000
Calling ibv_reg_dmabuf_mr(offset=0, size=8388608, addr=0x7416c5600000, fd=8) for QP #0
---------------------------------------------------------------------------------------
                    RDMA_Write Bidirectional BW Test
 Dual-port       : OFF          Device         : ionic_1
 Number of qps   : 4            Transport type : IB
 Connection type : RC           Using SRQ      : OFF
 PCIe relax order: ON           Lock-free      : OFF
 ibv_wr* API     : OFF          Using DDP      : OFF
 TX depth        : 128
 CQ Moderation   : 1
 CQE Poll Batch  : 16
 Mtu             : 4096[B]
 Link type       : Ethernet
 GID index       : 2
 Max inline data : 0[B]
 rdma_cm QPs     : OFF
 Use ROCm memory : ON
 Data ex. method : Ethernet
---------------------------------------------------------------------------------------
 local address: LID 0000 QPN 0x0800 PSN 0x8269b8 RKey 0x000106 VAddr 0x007416c5a00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0002 PSN 0xe9bd42 RKey 0x000106 VAddr 0x007416c5b00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0801 PSN 0xd86264 RKey 0x000106 VAddr 0x007416c5c00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 local address: LID 0000 QPN 0x0003 PSN 0xde5be3 RKey 0x000106 VAddr 0x007416c5d00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
 remote address: LID 0000 QPN 0x0800 PSN 0xe7c013 RKey 0x000106 VAddr 0x0073b224e00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0002 PSN 0x894401 RKey 0x000106 VAddr 0x0073b224f00000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 remote address: LID 0000 QPN 0x0801 PSN 0xe85bb7 RKey 0x000106 VAddr 0x0073b225000000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
 remote address: LID 0000 QPN 0x0003 PSN 0x30163a RKey 0x000106 VAddr 0x0073b225100000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
---------------------------------------------------------------------------------------
 #bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
 1048576    1392089          0.00               779.68               0.092945
---------------------------------------------------------------------------------------
deallocating GPU buffer 0x7416c5600000
```

6.2 MTU sweep using ib_write_bw

An MTU sweep test is used to validate basic NIC and PCIe functionality while measuring the average bandwidth of a specific RoCEv2 NIC. Running the test across multiple RoCEv2 MTU sizes demonstrates how larger MTU values achieve higher bandwidth. The test is conducted between NIC1 on server-1 (server) and NIC1 on server-2 (client), with traffic passing through the switch.

The logs shown below (server-side only) illustrate how the test is executed using a single NIC, while the accompanying chart summarizes results across all NICs.

The test is performed by first running the following command on the server side. For each test, only the MTU size changes.

```
./ib_write_bw -d ionic_1 -n 5000 -F -x 2 -s 1M -q 32 --report_gbits -p 5001 --ipv6 --ipv6-addr –sl
3 -m 4096
```

Afterward, in another terminal window, we ran the following command to start the client on server-2. The specified IPv6 address is the one assigned to the interface mapped to RDMA interface ionic_1.

```
./ib_write_bw -d ionic_1 -n 5000 -F -x 2 -s 1M -q 32 --report_gbits -p 5001 --ipv6 --ipv6-addr
fd00:2:9:1:1:1:0:2 --sl 3 -m 4096
```

6.2.1 RoCEv2 MTU size 4096

```
./ib_write_bw -d ionic_1 -n 5000 -F -x 2 -s 1M -q 32 --report_gbits -p 5001 --ipv6 --ipv6-addr
fd00:2:9:1:1:1:0:2 –-sl 3 -m 4096
---------------------------------------------------------------------------------------
                    RDMA_Write BW Test
 Dual-port       : OFF          Device         : ionic_1
 Number of qps   : 32           Transport type : IB
 Connection type : RC           Using SRQ      : OFF
 PCIe relax order: ON           Lock-free      : OFF
 ibv_wr* API     : OFF          Using DDP      : OFF
 TX depth        : 128
 CQ Moderation   : 1
 CQE Poll Batch  : Dynamic
 Mtu             : 4096[B]
```

```
 Link type       : Ethernet
 GID index       : 2
 Max inline data : 0[B]
 rdma_cm QPs     : OFF
 Use ROCm memory : OFF
 Data ex. method : Ethernet
---------------------------------------------------------------------------------------
 local address: LID 0000 QPN 0x0800 PSN 0x7f24fa RKey 0x00011a VAddr 0x007ed2bb18d000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
…
<snipped>
…
 remote address: LID 0000 QPN 0x0011 PSN 0x925589 RKey 0x00011a VAddr 0x0075f2dffc6000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
---------------------------------------------------------------------------------------
 #bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
 1048576    160000           389.87             389.84               0.046473
---------------------------------------------------------------------------------------
```

6.2.2 RoCEv2 MTU size 2048

```
./ib_write_bw -d ionic_1 -n 5000 -F -x 2 -s 1M -q 32 --report_gbits -p 5001 --ipv6 --ipv6-addr
fd00:2:9:1:1:1:0:2 --sl 3 -m 2048
---------------------------------------------------------------------------------------
                    RDMA_Write BW Test
Dual-port       : OFF          Device         : ionic_1
Number of qps   : 32           Transport type : IB
Connection type : RC           Using SRQ      : OFF
PCIe relax order: ON           Lock-free      : OFF
ibv_wr* API     : OFF          Using DDP      : OFF
TX depth        : 128
CQ Moderation   : 1
CQE Poll Batch  : Dynamic
Mtu             : 2048[B]
Link type       : Ethernet
GID index       : 2
Max inline data : 0[B]
rdma_cm QPs     : OFF
Use ROCm memory : OFF
Data ex. method : Ethernet
---------------------------------------------------------------------------------------
local address: LID 0000 QPN 0x0800 PSN 0x95954a RKey 0x00011c VAddr 0x007bdcac18d000
GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
…
<snipped>
…
remote address: LID 0000 QPN 0x0011 PSN 0x5900fa RKey 0x00011c VAddr 0x00790878cff000
GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
---------------------------------------------------------------------------------------
#bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
1048576    160000           380.61             380.35               0.045342
---------------------------------------------------------------------------------------
```

6.2.3 RoCEv2 MTU size 1024

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
./ib_write_bw -d ionic_1 -n 5000 -F -x 2 -s 1M -q 32 --report_gbits -p 5001 --ipv6 --ipv6-addr
fd00:2:9:1:1:1:0:2 --sl 3 -m 1024
---------------------------------------------------------------------------------------
                    RDMA_Write BW Test
Dual-port       : OFF          Device         : ionic_1
Number of qps   : 32           Transport type : IB
Connection type : RC           Using SRQ      : OFF
PCIe relax order: ON           Lock-free      : OFF
ibv_wr* API     : OFF          Using DDP      : OFF
TX depth        : 128
CQ Moderation   : 1
CQE Poll Batch  : Dynamic
Mtu             : 1024[B]
Link type       : Ethernet
GID index       : 2
Max inline data : 0[B]
rdma_cm QPs     : OFF
Use ROCm memory : OFF
Data ex. method : Ethernet
---------------------------------------------------------------------------------------
local address: LID 0000 QPN 0x0800 PSN 0xf936a6 RKey 0x00011b VAddr 0x0072efb64c6000
GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
…
<snipped>
…
remote address: LID 0000 QPN 0x0011 PSN 0x3d5db8 RKey 0x00011b VAddr 0x0075e3a78ff000
GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
---------------------------------------------------------------------------------------
#bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
1048576    160000           340.57             336.49               0.040113
---------------------------------------------------------------------------------------
```

6.2.4 RoCEv2 MTU size 256

```
./ib_write_bw -d ionic_1 -n 5000 -F -x 2 -s 1M -q 32 --report_gbits -p 5001 --ipv6 --ipv6-addr --
sl 3 -m 256
---------------------------------------------------------------------------------------
                    RDMA_Write BW Test
 Dual-port       : OFF          Device         : ionic_1
 Number of qps   : 32           Transport type : IB
 Connection type : RC           Using SRQ      : OFF
 PCIe relax order: ON           Lock-free      : OFF
 ibv_wr* API     : OFF          Using DDP      : OFF
 CQ Moderation   : 1
 CQE Poll Batch  : Dynamic
 Mtu             : 256[B]
 Link type       : Ethernet
 GID index       : 2
 Max inline data : 0[B]
 rdma_cm QPs     : OFF
 Use ROCm memory : OFF
 Data ex. method : Ethernet
---------------------------------------------------------------------------------------
 local address: LID 0000 QPN 0x0800 PSN 0xa496ff RKey 0x000193 VAddr 0x007f855b6c6000
 GID: 253:00:00:02:00:09:00:01:00:01:00:01:00:00:00:02
 <snipped>
 remote address: LID 0000 QPN 0x080f PSN 0xa6149b RKey 0x000144 VAddr 0x0071f8dcfc6000
 GID: 253:00:00:02:00:09:00:01:00:01:00:02:00:00:00:02
```

```
-------------------------------------------------------------------------------
 #bytes     #iterations    BW peak[Gb/sec]    BW average[Gb/sec]   MsgRate[Mpps]
 1048576    160000           74.82              73.69                0.008784
```

7 RCCL operations tests

RCCL performance tests per algorithm	
4-node AllReduce	358.01 GB/s
4-node ReduceScatter	361.74 GB/s
4-node Broadcast	343.18 GB/s
4-node All-to-All	56.79 GB/s

RCCL tests, using mpirun, are used to measure collective communication performance within a single node (a single AMD MI300X server) and across multiple nodes (up to four AMD MI300X servers, in this design). These tests include different collective algorithms such as AllReduce, ReduceScatter, Broadcast, and All-to-All. The All-to-All collective is the most stressful collective algorithm because it includes sending traffic from every rank to every other rank in your GPU cluster (with *N* nodes, you have an *N* × (*N*-1) send/receive). The All-to-All collective test ensures that communication completes within reasonable times, indicating that there are no PFC storms or deadlocks and that ECN, PFC, and buffer thresholds are all configured effectively.

7.1 AllReduce collective

```
cse@slate1:~/benchmarking/mpi-tests$ cat hostfile.txt
slate1.ncse.io slots=8
slate2.ncse.io slots=8
slate3.ncse.io slots=8
slate4.ncse.io slots=8
cse@slate1:~/benchmarking/mpi-tests$ cat rccl-pensando-pollara.sh
#!/bin/bash -x

export INSTALL_DIR=$HOME/ompi_for_gpu
export UCX_DIR=$INSTALL_DIR/ucx
export OMPI_DIR=$INSTALL_DIR/ompi
export RCCL_HOME=/home/cse/benchmarking/rccl-tests
export BUILD_DIR=/tmp/ompi_for_gpu_build
export LD_LIBRARY_PATH=$OMPI_DIR/lib:$UCX_DIR/lib:/opt/rocm/lib
export PATH=$OMPI_DIR/bin:$UCX_DIR/bin:$PATH

echo "LD Path -> $LD_LIBRARY_PATH"

    /home/cse/ompi_for_gpu/ompi/bin/mpirun -np 32 -N 8 --hostfile hostfile.txt \
      -x PATH=$PATH \
      -x LD_LIBRARY_PATH=$LD_LIBRARY_PATH \
      -x NCCL_SOCKET_IFNAME=ens50f1 \
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
      -x NCCL_IB_GID_INDEX=2 \
      -x UCX_IB_GID_INDEX=2 \
      -x NCCL_IB_HCA=ionic_0,ionic_1,ionic_2,ionic_3,ionic_4,ionic_5,ionic_6,ionic_7 \
      -x
UCX_NET_DEVICES=enp9s0.1000,enp38s0.1000,enp70s0.1000,enp105s0.1000,enp137s0.1000,enp166s0.1000,enp198s0.1000,enp233s0.1000  \
      -x HIP_VISIBLE_DEVICES=0,1,2,3,4,5,6,7 \
      -x NCCL_IB_PCI_RELAXED_ORDERING=1 \
      -x HSA_DISABLE_CACHE=1 \
      -x HSA_FORCE_FINE_GRAIN_PCIE=1 \
      -x NCCL_IB_TIMEOUT=22 \
      -x NCCL_IB_DISABLE=0 \
      -x NCCL_IB_SL=3 \
      --bind-to numa \
      --mca pml ucx \
      --mca osc ucx \
      --mca spml ucx \
      --mca btl ^vader,openib \
      --mca btl_tcp_if_include ens50f1 \
      /home/cse/benchmarking/rccl-tests/build/all_reduce_perf -b 8 -e 16G -f 2 -i 0 -g 1

cse@slate1:~/benchmarking/mpi-tests$ ./rccl-pensando-pollara.sh
+ export INSTALL_DIR=/home/cse/ompi_for_gpu
+ INSTALL_DIR=/home/cse/ompi_for_gpu
+ export UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ export OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ export RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ export BUILD_DIR=/tmp/ompi_for_gpu_build
+ BUILD_DIR=/tmp/ompi_for_gpu_build
+ export LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ export
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
+
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
+ echo 'LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib'
LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ /home/cse/ompi_for_gpu/ompi/bin/mpirun -np 32 -N 8 --hostfile hostfile.txt -x
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin -x
LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib -x NCCL_SOCKET_IFNAME=ens50f1 -x
NCCL_IB_GID_INDEX=2 -x UCX_IB_GID_INDEX=2 -x NCCL_IB_HCA=ionic_0,ionic_1,ionic_2,ionic_3,ionic_4,ionic_5,ionic_6,ionic_7 -x
UCX_NET_DEVICES=enp9s0.1000,enp38s0.1000,enp70s0.1000,enp105s0.1000,enp137s0.1000,enp166s0.1000,enp198s0.1000,enp233s0.1000 -x
HIP_VISIBLE_DEVICES=0,1,2,3,4,5,6,7 -x NCCL_IB_PCI_RELAXED_ORDERING=1 -x HSA_DISABLE_CACHE=1 -x HSA_FORCE_FINE_GRAIN_PCIE=1 -x
NCCL_IB_TIMEOUT=22 -x NCCL_IB_DISABLE=0 -x NCCL_IB_SL=3 --bind-to numa --mca pml ucx --mca osc ucx --mca spml ucx --mca btl
'^vader,openib' --mca btl_tcp_if_include ens50f1 /home/cse/benchmarking/rccl-tests/build/all_reduce_perf -b 8 -e 16G -f 2 -i 0
-g 1
# Collective test starting: all_reduce_perf
# nThread 1 nGpus 1 minBytes 8 maxBytes 17179869184 step: 2(factor) warmup iters: 5 iters: 20 agg iters: 1 validation: 1
graph: 0
#
rccl-tests: Version develop:e1b8a3a
# Using devices
#  Rank  0 Group  0 Pid  44014 on     slate1 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  1 Group  0 Pid  44015 on     slate1 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank  2 Group  0 Pid  44016 on     slate1 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank  3 Group  0 Pid  44017 on     slate1 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank  4 Group  0 Pid  44018 on     slate1 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank  5 Group  0 Pid  44019 on     slate1 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank  6 Group  0 Pid  44020 on     slate1 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank  7 Group  0 Pid  44021 on     slate1 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank  8 Group  0 Pid  41792 on     slate2 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  9 Group  0 Pid  41793 on     slate2 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 10 Group  0 Pid  41794 on     slate2 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 11 Group  0 Pid  41796 on     slate2 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 12 Group  0 Pid  41798 on     slate2 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 13 Group  0 Pid  41795 on     slate2 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 14 Group  0 Pid  41797 on     slate2 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 15 Group  0 Pid  41799 on     slate2 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 16 Group  0 Pid  41296 on     slate3 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 17 Group  0 Pid  41297 on     slate3 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 18 Group  0 Pid  41299 on     slate3 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 19 Group  0 Pid  41301 on     slate3 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 20 Group  0 Pid  41298 on     slate3 device  4 [0000:85:00] AMD Instinct MI300X
```

```
#  Rank 21 Group  0 Pid  41300 on     slate3 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 22 Group  0 Pid  41302 on     slate3 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 23 Group  0 Pid  41303 on     slate3 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 24 Group  0 Pid  41096 on     slate4 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 25 Group  0 Pid  41098 on     slate4 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 26 Group  0 Pid  41097 on     slate4 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 27 Group  0 Pid  41101 on     slate4 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 28 Group  0 Pid  41099 on     slate4 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 29 Group  0 Pid  41102 on     slate4 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 30 Group  0 Pid  41100 on     slate4 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 31 Group  0 Pid  41103 on     slate4 device  7 [0000:e5:00] AMD Instinct MI300X
#
#                                                              out-of-place                       in-place
#       size         count      type   redop    root     time   algbw   busbw #wrong     time   algbw   busbw #wrong
#        (B)    (elements)                               (us)  (GB/s)  (GB/s)            (us)  (GB/s)  (GB/s)
           8             2     float     sum      -1    61.50    0.00    0.00      0    59.81    0.00    0.00      0
          16             4     float     sum      -1    59.07    0.00    0.00      0    58.81    0.00    0.00      0
          32             8     float     sum      -1    58.87    0.00    0.00      0    59.12    0.00    0.00      0
          64            16     float     sum      -1    58.28    0.00    0.00      0    62.89    0.00    0.00      0
         128            32     float     sum      -1    59.03    0.00    0.00      0    59.24    0.00    0.00      0
         256            64     float     sum      -1    59.70    0.00    0.01      0    59.61    0.00    0.01      0
         512           128     float     sum      -1    60.09    0.01    0.02      0    59.94    0.01    0.02      0
        1024           256     float     sum      -1    60.65    0.02    0.03      0    61.54    0.02    0.03      0
        2048           512     float     sum      -1    66.09    0.03    0.06      0    65.68    0.03    0.06      0
        4096          1024     float     sum      -1    67.36    0.06    0.12      0    67.15    0.06    0.12      0
        8192          2048     float     sum      -1    68.97    0.12    0.23      0    68.15    0.12    0.23      0
       16384          4096     float     sum      -1    71.25    0.23    0.45      0    70.65    0.23    0.45      0
       32768          8192     float     sum      -1    73.24    0.45    0.87      0    73.05    0.45    0.87      0
       65536         16384     float     sum      -1    81.78    0.80    1.55      0    81.07    0.81    1.57      0
      131072         32768     float     sum      -1    88.17    1.49    2.88      0    85.67    1.53    2.96      0
      262144         65536     float     sum      -1    100.0    2.62    5.08      0    100.6    2.60    5.05      0
      524288        131072     float     sum      -1    128.5    4.08    7.90      0    130.7    4.01    7.77      0
     1048576        262144     float     sum      -1    242.5    4.32    8.38      0    237.2    4.42    8.57      0
     2097152        524288     float     sum      -1    204.2   10.27   19.90      0    204.3   10.26   19.89      0
     4194304       1048576     float     sum      -1    242.1   17.32   33.56      0    244.8   17.14   33.20      0
     8388608       2097152     float     sum      -1    333.4   25.16   48.74      0    332.2   25.25   48.92      0
    16777216       4194304     float     sum      -1    460.5   36.43   70.59      0    462.5   36.27   70.28      0
    33554432       8388608     float     sum      -1    790.5   42.45   82.24      0    786.5   42.66   82.66      0
    67108864      16777216     float     sum      -1    915.9   73.27  141.96      0    917.1   73.18  141.78      0
   134217728      33554432     float     sum      -1   1593.4   84.23  163.20      0   1588.2   84.51  163.74      0
   268435456      67108864     float     sum      -1   2919.2   91.95  178.16      0   2919.9   91.93  178.12      0
   536870912     134217728     float     sum      -1   3071.8  174.77  338.62      0   3064.2  175.21  339.47      0
  1073741824     268435456     float     sum      -1   5888.7  182.34  353.28      0   5889.8  182.31  353.22      0
  2147483648     536870912     float     sum      -1    11670  184.02  356.53      0    11672  183.99  356.48      0
  4294967296    1073741824     float     sum      -1    23302  184.32  357.12      0    23291  184.40  357.28      0
  8589934592    2147483648     float     sum      -1    46613  184.28  357.05      0    46582  184.40  357.28      0
 17179869184    4294967296     float     sum      -1    92953  184.82  358.09      0    93054  184.62  357.71      0
# Errors with asterisks indicate errors that have exceeded the maximum threshold.
# Out of bounds values : 0 OK
# Avg bus bandwidth    : 90.2246
#
# Collective test concluded: all_reduce_perf
```

7.2 ReduceScatter collective

```
cse@slate1:~/benchmarking/mpi-tests$ ./rccl-pensando-pollara.sh
+ export INSTALL_DIR=/home/cse/ompi_for_gpu
+ INSTALL_DIR=/home/cse/ompi_for_gpu
+ export UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ export OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ export RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ export BUILD_DIR=/tmp/ompi_for_gpu_build
+ BUILD_DIR=/tmp/ompi_for_gpu_build
+ export LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ export
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
+
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
+ echo 'LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib'
LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ /home/cse/ompi_for_gpu/ompi/bin/mpirun -np 32 -N 8 --hostfile hostfile.txt -x
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin -x
LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib -x NCCL_SOCKET_IFNAME=ens50f1 -x
NCCL_IB_GID_INDEX=2 -x UCX_IB_GID_INDEX=2 -x NCCL_IB_HCA=ionic_0,ionic_1,ionic_2,ionic_3,ionic_4,ionic_5,ionic_6,ionic_7 -x
UCX_NET_DEVICES=enp9s0.1000,enp38s0.1000,enp70s0.1000,enp105s0.1000,enp137s0.1000,enp166s0.1000,enp198s0.1000,enp233s0.1000 -x
HIP_VISIBLE_DEVICES=0,1,2,3,4,5,6,7 -x NCCL_IB_PCI_RELAXED_ORDERING=1 -x HSA_DISABLE_CACHE=1 -x HSA_FORCE_FINE_GRAIN_PCIE=1 -x
NCCL_IB_TIMEOUT=22 -x NCCL_IB_DISABLE=0 -x NCCL_IB_SL=3 --bind-to numa --mca pml ucx --mca osc ucx --mca spml ucx --mca btl
'^vader,openib' --mca btl_tcp_if_include ens50f1 /home/cse/benchmarking/rccl-tests/build/reduce_scatter_perf -b 8 -e 16G -f 2
-i 0 -g 1
# Collective test starting: reduce_scatter_perf
# nThread 1 nGpus 1 minBytes 8 maxBytes 17179869184 step: 2(factor) warmup iters: 5 iters: 20 agg iters: 1 validation: 1
graph: 0
#
rccl-tests: Version develop:e1b8a3a
# Using devices
#  Rank  0 Group  0 Pid  44475 on     slate1 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  1 Group  0 Pid  44476 on     slate1 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank  2 Group  0 Pid  44477 on     slate1 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank  3 Group  0 Pid  44478 on     slate1 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank  4 Group  0 Pid  44479 on     slate1 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank  5 Group  0 Pid  44480 on     slate1 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank  6 Group  0 Pid  44481 on     slate1 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank  7 Group  0 Pid  44482 on     slate1 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank  8 Group  0 Pid  42287 on     slate2 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  9 Group  0 Pid  42288 on     slate2 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 10 Group  0 Pid  42289 on     slate2 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 11 Group  0 Pid  42291 on     slate2 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 12 Group  0 Pid  42292 on     slate2 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 13 Group  0 Pid  42290 on     slate2 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 14 Group  0 Pid  42293 on     slate2 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 15 Group  0 Pid  42294 on     slate2 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 16 Group  0 Pid  41859 on     slate3 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 17 Group  0 Pid  41860 on     slate3 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 18 Group  0 Pid  41861 on     slate3 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 19 Group  0 Pid  41862 on     slate3 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 20 Group  0 Pid  41863 on     slate3 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 21 Group  0 Pid  41864 on     slate3 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 22 Group  0 Pid  41865 on     slate3 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 23 Group  0 Pid  41866 on     slate3 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 24 Group  0 Pid  41640 on     slate4 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 25 Group  0 Pid  41641 on     slate4 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 26 Group  0 Pid  41642 on     slate4 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 27 Group  0 Pid  41643 on     slate4 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 28 Group  0 Pid  41645 on     slate4 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 29 Group  0 Pid  41644 on     slate4 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 30 Group  0 Pid  41646 on     slate4 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 31 Group  0 Pid  41647 on     slate4 device  7 [0000:e5:00] AMD Instinct MI300X
#
#                                                              out-of-place                       in-place
#       size         count      type   redop    root     time   algbw   busbw #wrong     time   algbw   busbw #wrong
#        (B)    (elements)                               (us)  (GB/s)  (GB/s)            (us)  (GB/s)  (GB/s)
           0             0     float     sum      -1     0.17    0.00    0.00      0     0.09    0.00    0.00      0
           0             0     float     sum      -1     0.09    0.00    0.00      0     0.09    0.00    0.00      0
           0             0     float     sum      -1     0.09    0.00    0.00      0     0.09    0.00    0.00      0
           0             0     float     sum      -1     0.09    0.00    0.00      0     0.09    0.00    0.00      0
           0             0     float     sum      -1     0.09    0.00    0.00      0     0.09    0.00    0.00      0
           0             0     float     sum      -1     0.09    0.00    0.00      0     0.12    0.00    0.00      0
         512             4     float     sum      -1    84.85    0.01    0.01      0    84.18    0.01    0.01      0
        1024             8     float     sum      -1    83.83    0.01    0.01      0    83.96    0.01    0.01      0
        2048            16     float     sum      -1    84.19    0.02    0.02      0    84.44    0.02    0.02      0
        4096            32     float     sum      -1    84.04    0.05    0.05      0    83.69    0.05    0.05      0
        8192            64     float     sum      -1    84.28    0.10    0.09      0    84.39    0.10    0.09      0
       16384           128     float     sum      -1    85.30    0.19    0.19      0    85.33    0.19    0.19      0
       32768           256     float     sum      -1    86.62    0.38    0.37      0    87.12    0.38    0.36      0
       65536           512     float     sum      -1    88.13    0.74    0.72      0    88.41    0.74    0.72      0
      131072          1024     float     sum      -1    89.79    1.46    1.41      0    88.85    1.48    1.43      0
      262144          2048     float     sum      -1    89.62    2.93    2.83      0    88.83    2.95    2.86      0
      524288          4096     float     sum      -1    110.3    4.75    4.60      0    108.1    4.85    4.70      0
     1048576          8192     float     sum      -1    169.5    6.18    5.99      0    167.0    6.28    6.08      0
     2097152         16384     float     sum      -1    300.1    6.99    6.77      0    291.6    7.19    6.97      0
     4194304         32768     float     sum      -1    571.3    7.34    7.11      0    559.7    7.49    7.26      0
     8388608         65536     float     sum      -1    386.1   21.73   21.05      0    387.4   21.65   20.98      0
    16777216        131072     float     sum      -1    398.8   42.07   40.76      0    393.5   42.63   41.30      0
    33554432        262144     float     sum      -1    421.9   79.53   77.05      0    422.1   79.49   77.01      0
    67108864        524288     float     sum      -1    531.3  126.32  122.37      0    530.5  126.50  122.55      0
```

```
   134217728       1048576     float     sum      -1   1154.1  116.29  112.66      0   1120.0  119.84  116.10      0
   268435456       2097152     float     sum      -1   1209.8  221.88  214.95      0   1206.6  222.48  215.53      0
   536870912       4194304     float     sum      -1   1531.8  350.49  339.54      0   1524.4  352.19  341.19      0
  1073741824       8388608     float     sum      -1   2957.8  363.02  351.68      0   2824.6  380.14  368.26      0
  2147483648      16777216     float     sum      -1   5839.8  367.73  356.24      0   5585.9  384.45  372.43      0
  4294967296      33554432     float     sum      -1    11642  368.93  357.40      0    11179  384.19  372.18      0
  8589934592      67108864     float     sum      -1    23147  371.10  359.51      0    22545  381.02  369.11      0
 17179869184     134217728     float     sum      -1    46008  373.41  361.74      0    46014  373.36  361.69      0
# Errors with asterisks indicate errors that have exceeded the maximum threshold.
# Out of bounds values : 0 OK
# Avg bus bandwidth    : 86.7841
#
# Collective test concluded: reduce_scatter_perf
```

7.3 Broadcast collective

```
cse@slate1:~/benchmarking/mpi-tests$ ./rccl-pensando-pollara.sh
+ export INSTALL_DIR=/home/cse/ompi_for_gpu
+ INSTALL_DIR=/home/cse/ompi_for_gpu
+ export UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ export OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ export RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ export BUILD_DIR=/tmp/ompi_for_gpu_build
+ BUILD_DIR=/tmp/ompi_for_gpu_build
+ export LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ export
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
+
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
+ echo 'LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib'
LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ /home/cse/ompi_for_gpu/ompi/bin/mpirun -np 32 -N 8 --hostfile hostfile.txt -x
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin -x
LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib -x NCCL_SOCKET_IFNAME=ens50f1 -x
NCCL_IB_GID_INDEX=2 -x UCX_IB_GID_INDEX=2 -x NCCL_IB_HCA=ionic_0,ionic_1,ionic_2,ionic_3,ionic_4,ionic_5,ionic_6,ionic_7 -x
UCX_NET_DEVICES=enp9s0.1000,enp38s0.1000,enp70s0.1000,enp105s0.1000,enp137s0.1000,enp166s0.1000,enp198s0.1000,enp233s0.1000 -x
HIP_VISIBLE_DEVICES=0,1,2,3,4,5,6,7 -x NCCL_IB_PCI_RELAXED_ORDERING=1 -x HSA_DISABLE_CACHE=1 -x HSA_FORCE_FINE_GRAIN_PCIE=1 -x
NCCL_IB_TIMEOUT=22 -x NCCL_IB_DISABLE=0 -x NCCL_IB_SL=3 --bind-to numa --mca pml ucx --mca osc ucx --mca spml ucx --mca btl
'^vader,openib' --mca btl_tcp_if_include ens50f1 /home/cse/benchmarking/rccl-tests/build/broadcast_perf -b 8 -e 16G -f 2 -i 0
-g 1
# Collective test starting: broadcast_perf
# nThread 1 nGpus 1 minBytes 8 maxBytes 17179869184 step: 2(factor) warmup iters: 5 iters: 20 agg iters: 1 validation: 1
graph: 0
#
rccl-tests: Version develop:e1b8a3a
# Using devices
#  Rank  0 Group  0 Pid  44773 on     slate1 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  1 Group  0 Pid  44774 on     slate1 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank  2 Group  0 Pid  44775 on     slate1 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank  3 Group  0 Pid  44776 on     slate1 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank  4 Group  0 Pid  44777 on     slate1 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank  5 Group  0 Pid  44778 on     slate1 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank  6 Group  0 Pid  44779 on     slate1 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank  7 Group  0 Pid  44780 on     slate1 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank  8 Group  0 Pid  42722 on     slate2 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  9 Group  0 Pid  42724 on     slate2 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 10 Group  0 Pid  42723 on     slate2 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 11 Group  0 Pid  42725 on     slate2 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 12 Group  0 Pid  42727 on     slate2 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 13 Group  0 Pid  42726 on     slate2 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 14 Group  0 Pid  42728 on     slate2 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 15 Group  0 Pid  42729 on     slate2 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 16 Group  0 Pid  42256 on     slate3 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 17 Group  0 Pid  42257 on     slate3 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 18 Group  0 Pid  42258 on     slate3 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 19 Group  0 Pid  42259 on     slate3 device  3 [0000:65:00] AMD Instinct MI300X
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
#  Rank 20 Group  0 Pid  42260 on     slate3 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 21 Group  0 Pid  42261 on     slate3 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 22 Group  0 Pid  42262 on     slate3 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 23 Group  0 Pid  42263 on     slate3 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 24 Group  0 Pid  42039 on     slate4 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 25 Group  0 Pid  42040 on     slate4 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 26 Group  0 Pid  42041 on     slate4 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 27 Group  0 Pid  42046 on     slate4 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 28 Group  0 Pid  42043 on     slate4 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 29 Group  0 Pid  42042 on     slate4 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 30 Group  0 Pid  42044 on     slate4 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 31 Group  0 Pid  42045 on     slate4 device  7 [0000:e5:00] AMD Instinct MI300X
#
#                                                              out-of-place                       in-place
#       size         count      type   redop    root     time   algbw   busbw #wrong     time   algbw   busbw #wrong
#        (B)    (elements)                               (us)  (GB/s)  (GB/s)            (us)  (GB/s)  (GB/s)
           8             2     float    none       0    19.94    0.00    0.00      0    15.11    0.00    0.00      0
          16             4     float    none       0    16.33    0.00    0.00      0    15.00    0.00    0.00      0
          32             8     float    none       0    15.31    0.00    0.00      0    15.18    0.00    0.00      0
          64            16     float    none       0    15.56    0.00    0.00      0    18.67    0.00    0.00      0
         128            32     float    none       0    15.47    0.01    0.01      0    18.31    0.01    0.01      0
         256            64     float    none       0    14.54    0.02    0.02      0    14.91    0.02    0.02      0
         512           128     float    none       0    14.64    0.03    0.03      0    14.83    0.03    0.03      0
        1024           256     float    none       0    14.92    0.07    0.07      0    15.21    0.07    0.07      0
        2048           512     float    none       0    14.96    0.14    0.14      0    15.19    0.13    0.13      0
        4096          1024     float    none       0    16.33    0.25    0.25      0    16.31    0.25    0.25      0
        8192          2048     float    none       0    17.75    0.46    0.46      0    17.63    0.46    0.46      0
       16384          4096     float    none       0    21.19    0.77    0.77      0    20.86    0.79    0.79      0
       32768          8192     float    none       0    27.49    1.19    1.19      0    27.33    1.20    1.20      0
       65536         16384     float    none       0    39.86    1.64    1.64      0    39.17    1.67    1.67      0
      131072         32768     float    none       0    80.72    1.62    1.62      0    79.70    1.64    1.64      0
      262144         65536     float    none       0    120.9    2.17    2.17      0    120.3    2.18    2.18      0
      524288        131072     float    none       0    123.2    4.25    4.25      0    123.0    4.26    4.26      0
     1048576        262144     float    none       0    146.9    7.14    7.14      0    144.2    7.27    7.27      0
     2097152        524288     float    none       0    211.4    9.92    9.92      0    211.4    9.92    9.92      0
     4194304       1048576     float    none       0    278.7   15.05   15.05      0    276.5   15.17   15.17      0
     8388608       2097152     float    none       0    317.0   26.46   26.46      0    317.4   26.43   26.43      0
    16777216       4194304     float    none       0    403.3   41.60   41.60      0    408.0   41.12   41.12      0
    33554432       8388608     float    none       0    520.2   64.50   64.50      0    523.9   64.04   64.04      0
    67108864      16777216     float    none       0    679.0   98.83   98.83      0    677.3   99.09   99.09      0
   134217728      33554432     float    none       0   1087.7  123.40  123.40      0   1086.0  123.59  123.59      0
   268435456      67108864     float    none       0   1506.0  178.24  178.24      0   1506.4  178.20  178.20      0
   536870912     134217728     float    none       0   2807.3  191.24  191.24      0   2798.6  191.83  191.83      0
  1073741824     268435456     float    none       0   4330.7  247.94  247.94      0   4335.7  247.65  247.65      0
  2147483648     536870912     float    none       0   8455.9  253.96  253.96      0   8465.8  253.67  253.67      0
  4294967296    1073741824     float    none       0    14419  297.88  297.88      0    14417  297.90  297.90      0
  8589934592    2147483648     float    none       0    26287  326.78  326.78      0    26263  327.08  327.08      0
 17179869184    4294967296     float    none       0    50060  343.18  343.18      0    49941  344.00  344.00      0
# Errors with asterisks indicate errors that have exceeded the maximum threshold.
# Out of bounds values : 0 OK
# Avg bus bandwidth    : 69.9757
#
# Collective test concluded: broadcast_perf
```

7.4 All-to-All collective

```
cse@slate1:~/benchmarking/mpi-tests$ ./rccl-pensando-pollara-alltoall.sh
+ export INSTALL_DIR=/home/cse/ompi_for_gpu
+ INSTALL_DIR=/home/cse/ompi_for_gpu
+ export UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ UCX_DIR=/home/cse/ompi_for_gpu/ucx
+ export OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ OMPI_DIR=/home/cse/ompi_for_gpu/ompi
+ export RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ RCCL_HOME=/home/cse/benchmarking/rccl-tests
+ export BUILD_DIR=/tmp/ompi_for_gpu_build
+ BUILD_DIR=/tmp/ompi_for_gpu_build
+ export LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ export
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
```

```
+
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin
+ echo 'LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib'
LD Path -> /home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib
+ /home/cse/ompi_for_gpu/ompi/bin/mpirun -np 32 -N 8 --hostfile hostfile.txt -x
PATH=/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/b
in:/usr/games:/usr/local/games:/snap/bin:/opt/rocm-6.4.3/bin:/home/cse/ompi_for_gpu/ompi/bin:/home/cse/ompi_for_gpu/ucx/bin -x
LD_LIBRARY_PATH=/home/cse/ompi_for_gpu/ompi/lib:/home/cse/ompi_for_gpu/ucx/lib:/opt/rocm/lib -x NCCL_SOCKET_IFNAME=ens50f1 -x
NCCL_IB_GID_INDEX=2 -x UCX_IB_GID_INDEX=2 -x NCCL_IB_HCA=ionic_0,ionic_1,ionic_2,ionic_3,ionic_4,ionic_5,ionic_6,ionic_7 -x
UCX_NET_DEVICES=enp9s0.1000,enp38s0.1000,enp70s0.1000,enp105s0.1000,enp137s0.1000,enp166s0.1000,enp198s0.1000,enp233s0.1000 -x
HIP_VISIBLE_DEVICES=0,1,2,3,4,5,6,7 -x NCCL_IB_PCI_RELAXED_ORDERING=1 -x HSA_DISABLE_CACHE=1 -x HSA_FORCE_FINE_GRAIN_PCIE=1 -x
NCCL_IB_TIMEOUT=22 -x NCCL_IB_DISABLE=0 -x NCCL_PXN_DISABLE=0 -x NCCL_IB_SL=3 --bind-to numa --mca pml ucx --mca osc ucx --mca
spml ucx --mca btl '^vader,openib' --mca btl_tcp_if_include ens50f1 /home/cse/benchmarking/rccl-tests/build/alltoall_perf -b 8
-e 16G -f 2 -i 0 -g 1
# Collective test starting: alltoall_perf
# nThread 1 nGpus 1 minBytes 8 maxBytes 17179869184 step: 2(factor) warmup iters: 5 iters: 20 agg iters: 1 validation: 1
graph: 0
#
rccl-tests: Version develop:e1b8a3a
# Using devices
#  Rank  0 Group  0 Pid  45252 on     slate1 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  1 Group  0 Pid  45253 on     slate1 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank  2 Group  0 Pid  45254 on     slate1 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank  3 Group  0 Pid  45255 on     slate1 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank  4 Group  0 Pid  45256 on     slate1 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank  5 Group  0 Pid  45257 on     slate1 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank  6 Group  0 Pid  45258 on     slate1 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank  7 Group  0 Pid  45259 on     slate1 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank  8 Group  0 Pid  43272 on     slate2 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank  9 Group  0 Pid  43273 on     slate2 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 10 Group  0 Pid  43275 on     slate2 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 11 Group  0 Pid  43274 on     slate2 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 12 Group  0 Pid  43276 on     slate2 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 13 Group  0 Pid  43277 on     slate2 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 14 Group  0 Pid  43278 on     slate2 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 15 Group  0 Pid  43279 on     slate2 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 16 Group  0 Pid  42775 on     slate3 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 17 Group  0 Pid  42776 on     slate3 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 18 Group  0 Pid  42777 on     slate3 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 19 Group  0 Pid  42778 on     slate3 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 20 Group  0 Pid  42781 on     slate3 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 21 Group  0 Pid  42779 on     slate3 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 22 Group  0 Pid  42780 on     slate3 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 23 Group  0 Pid  42782 on     slate3 device  7 [0000:e5:00] AMD Instinct MI300X
#  Rank 24 Group  0 Pid  42517 on     slate4 device  0 [0000:05:00] AMD Instinct MI300X
#  Rank 25 Group  0 Pid  42518 on     slate4 device  1 [0000:29:00] AMD Instinct MI300X
#  Rank 26 Group  0 Pid  42521 on     slate4 device  2 [0000:49:00] AMD Instinct MI300X
#  Rank 27 Group  0 Pid  42520 on     slate4 device  3 [0000:65:00] AMD Instinct MI300X
#  Rank 28 Group  0 Pid  42522 on     slate4 device  4 [0000:85:00] AMD Instinct MI300X
#  Rank 29 Group  0 Pid  42519 on     slate4 device  5 [0000:a9:00] AMD Instinct MI300X
#  Rank 30 Group  0 Pid  42524 on     slate4 device  6 [0000:c9:00] AMD Instinct MI300X
#  Rank 31 Group  0 Pid  42523 on     slate4 device  7 [0000:e5:00] AMD Instinct MI300X
#
#                                                              out-of-place                       in-place
#       size         count      type   redop    root     time   algbw   busbw #wrong     time   algbw   busbw #wrong
#        (B)    (elements)                               (us)  (GB/s)  (GB/s)            (us)  (GB/s)  (GB/s)
           0             0     float    none      -1     0.16    0.00    0.00      0     0.09    0.00    0.00    N/A
           0             0     float    none      -1     0.06    0.00    0.00      0     0.05    0.00    0.00    N/A
           0             0     float    none      -1     0.05    0.00    0.00      0     0.05    0.00    0.00    N/A
           0             0     float    none      -1     0.05    0.00    0.00      0     0.05    0.00    0.00    N/A
           0             0     float    none      -1     0.05    0.00    0.00      0     0.06    0.00    0.00    N/A
           0             0     float    none      -1     0.05    0.00    0.00      0     0.06    0.00    0.00    N/A
         512             4     float    none      -1    83.82    0.01    0.01      0    86.33    0.01    0.01    N/A
        1024             8     float    none      -1    84.71    0.01    0.01      0    83.46    0.01    0.01    N/A
        2048            16     float    none      -1    84.23    0.02    0.02      0    85.14    0.02    0.02    N/A
        4096            32     float    none      -1    85.74    0.05    0.05      0    84.22    0.05    0.05    N/A
        8192            64     float    none      -1    84.08    0.10    0.09      0    83.45    0.10    0.10    N/A
       16384           128     float    none      -1    84.93    0.19    0.19      0    84.69    0.19    0.19    N/A
       32768           256     float    none      -1    85.25    0.38    0.37      0    84.84    0.39    0.37    N/A
       65536           512     float    none      -1    84.31    0.78    0.75      0    83.98    0.78    0.76    N/A
      131072          1024     float    none      -1    85.50    1.53    1.49      0    86.19    1.52    1.47    N/A
      262144          2048     float    none      -1    85.39    3.07    2.97      0    86.78    3.02    2.93    N/A
      524288          4096     float    none      -1    85.78    6.11    5.92      0    88.49    5.93    5.74    N/A
     1048576          8192     float    none      -1    86.28   12.15   11.77      0    88.44   11.86   11.49    N/A
     2097152         16384     float    none      -1    136.0   15.42   14.94      0    136.0   15.42   14.93    N/A
     4194304         32768     float    none      -1    185.7   22.59   21.88      0    178.8   23.45   22.72    N/A
     8388608         65536     float    none      -1    286.2   29.31   28.39      0    285.2   29.42   28.50    N/A
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms

```
    16777216        131072     float    none      -1    486.2   34.50   33.43      0    497.6   33.71   32.66    N/A
    33554432        262144     float    none      -1    638.9   52.52   50.88      0    631.4   53.15   51.48    N/A
    67108864        524288     float    none      -1   1169.5   57.38   55.59      0   1183.6   56.70   54.93    N/A
   134217728       1048576     float    none      -1   2288.6   58.65   56.81      0   2284.4   58.75   56.92    N/A
   268435456       2097152     float    none      -1   4560.7   58.86   57.02      0   4548.9   59.01   57.17    N/A
   536870912       4194304     float    none      -1   9055.4   59.29   57.43      0   9051.5   59.31   57.46    N/A
  1073741824       8388608     float    none      -1    19094   56.23   54.48      0    18956   56.64   54.87    N/A
  2147483648      16777216     float    none      -1    37273   57.61   55.81      0    37449   57.34   55.55    N/A
  4294967296      33554432     float    none      -1    73593   58.36   56.54      0    73249   58.64   56.80    N/A
  8589934592      67108864     float    none      -1   145832   58.90   57.06      0   146165   58.77   56.93    N/A
 17179869184     134217728     float    none      -1   293052   58.62   56.79      0   292094   58.82   56.98    N/A
# Errors with asterisks indicate errors that have exceeded the maximum threshold.
# Out of bounds values : 0 OK
# Avg bus bandwidth    : 21.2769
#
# Collective test concluded: alltoall_perf
```


© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms