

AsterNOS Data Center AsterFabric Matrix Architecture Test Report

Version 1.0

Contents

1 Report Overview	1
1.1 Test Content.....	1
1.2 Test Results Overview	1
1.3 Test Summary	2
2 Test Environment.....	6
2.1 Test Topology.....	6
2.1.1 Baseline Test Topology	6
2.1.2 Cross-Switch Test Topology.....	6
2.1.3 AsterFabric ROFT Test Topology	7
2.1.4 AsterFabric Matrix Test Topology	7
2.2 Hardware and Software Environment.....	8
2.2.1 Switches	8
2.2.2 Compute Nodes	8
2.2.3 Optical Modules and Cables.....	9
3 E2E Forwarding Performance Test.....	10
3.1 NIC Back-to-Back Forwarding Bandwidth Test.....	10
3.2 Cross-Switch Forwarding Bandwidth Test.....	11
3.3 AsterFabric ROFT Forwarding Bandwidth Test	12
3.4 AsterFabric Matrix Forwarding Bandwidth Test	13
4 Collective Communication Performance Test	15
4.1 Cross-Switch CCL Test	15
4.1.1 Cross-Switch All-Reduce Test	15
4.1.2 Cross-Switch All-to-All Test	16
4.2 AsterFabric ROFT CCL Test.....	17
4.2.1 AsterFabric ROFT All-Reduce Test.....	17
4.2.2 AsterFabric ROFT All-to-All Test.....	18
4.3 AsterFabric Matrix CCL Test.....	19
4.3.1 AsterFabric Matrix All-Reduce Test.....	19
4.3.2 AsterFabric Matrix All-to-All Test.....	20
4.4 AsterFabric Matrix CCL Tuning Test.....	21
4.4.1 AsterFabric Matrix All-Reduce Tuning Test.....	21
4.4.2 AsterFabric Matrix All-to-All Tuning Test.....	22
5 Model Inference Performance Test.....	24
5.1 AsterFabric ROFT DeepSeek Model Inference Test	24
5.2 AsterFabric Matrix DeepSeek Model Inference Test	25

1 Report Overview

1.1 Test Content

This report describes in detail the performance testing of the Asterfusion CX-N ultra-low-latency switches from Asterfusion Data Technologies Co., Ltd. on the AsterFabric Matrix (Orthogonal Dual-Plane) architecture, with the aim of verifying whether this network architecture can deliver network performance comparable to that of a ROFT network. The test content mainly includes:

1. E2E forwarding performance test: uses the RDMA performance test toolset to measure the forwarding performance of RoCE traffic under each network architecture built with the switches under test.
2. CCL collective communication library performance test: uses test tools to verify the collective communication performance under each network architecture built with the switches under test.
3. Model inference performance test: deploys an inference model to verify the performance under each network architecture built with the switches under test.

1.2 Test Results Overview

Table 1 Test Results Overview

Test Type	Test Item	Test Result	Remarks
E2E Forwarding Performance	NIC Back-to-Back Forwarding Bandwidth Test	390.62 Gb/s	NIC in bond 400G mode
	Cross-Switch Forwarding Bandwidth Test	392.05 Gb/s	NIC in single-port 400G mode
	AsterFabric ROFT Forwarding Bandwidth Test	392.06 Gb/s	NIC in single-port 400G mode
	AsterFabric Matrix Forwarding Bandwidth Test	390.63 Gb/s	NIC in bond 400G mode
Collective Comm. Performance	Cross-Switch All-Reduce Test	155.66 GB/s	Pre-tuning parameters used
	Cross-Switch All-to-All Test	29.58 GB/s	
	AsterFabric ROFT All-Reduce Test	152.93 GB/s	
	AsterFabric ROFT All-to-All Test	29.67 GB/s	
	AsterFabric Matrix All-Reduce Test	153.01 GB/s	
	AsterFabric Matrix All-to-All Test	29.69 GB/s	
	AsterFabric Matrix All-Reduce Tuning Test	162.78 GB/s	Post-tuning parameters used
	AsterFabric Matrix All-to-All Tuning Test	30.11 GB/s	

Test Type	Test Item	Test Result	Remarks
Model Inference Performance	AsterFabric ROFT DeepSeek Model Inference Test	Total Throughput ^[i] : 10,765.62 token/s TTFT ^[ii] : 2,403.08 ms TPOT ^[iii] : 19.14 ms	DeepSeek-V4-Flash model 2P2D deployment Concurrency 64
	AsterFabric Matrix DeepSeek Model Inference Test	Total Throughput: 13,101.82 token/s TTFT: 1,215.97 ms TPOT: 17.34 ms	

1.3 Test Summary

1. In the E2E forwarding performance test, the ROFT network measured 392.06 Gb/s, essentially the same as the cross-switch reference baseline of 392.05 Gb/s; the Matrix network measured 390.63 Gb/s, essentially the same as the back-to-back reference baseline of 390.62 Gb/s. The results show that both networks achieved forwarding bandwidth close to their reference baselines under their respective test conditions. Because the two networks use different NIC operating modes and different reference baselines, these results are not intended for directly comparing the relative performance of the two architectures.

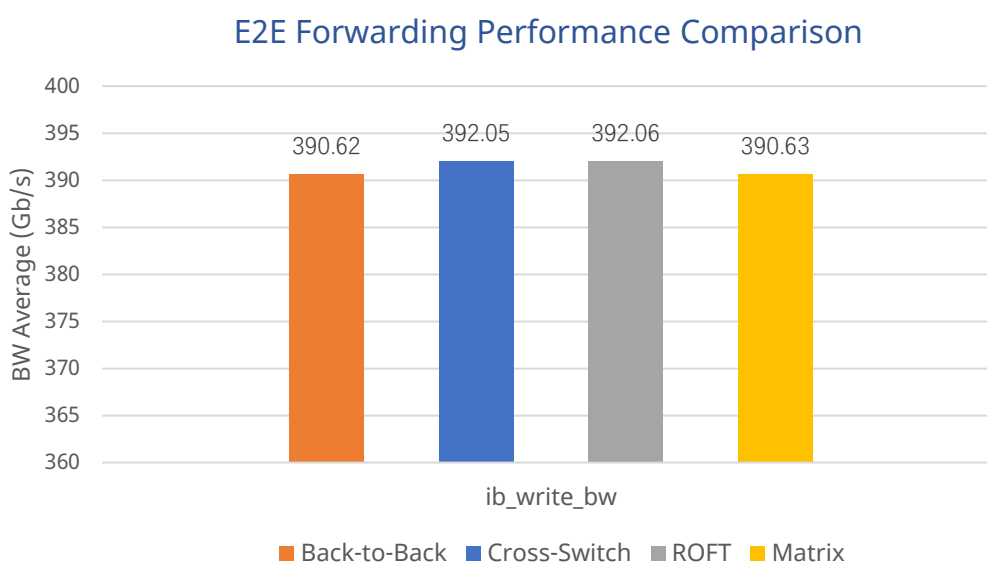


Figure 1 E2E Forwarding Performance Comparison

2. In the CCL performance test, with identical test parameters, the bus bandwidth at an 8G message size was taken as the test result. As shown in the figure below, the All-Reduce and All-to-All results of the Matrix network are essentially the same as those of ROFT, indicating that Matrix can deliver performance comparable to ROFT under the collective communication workloads covered in this test.

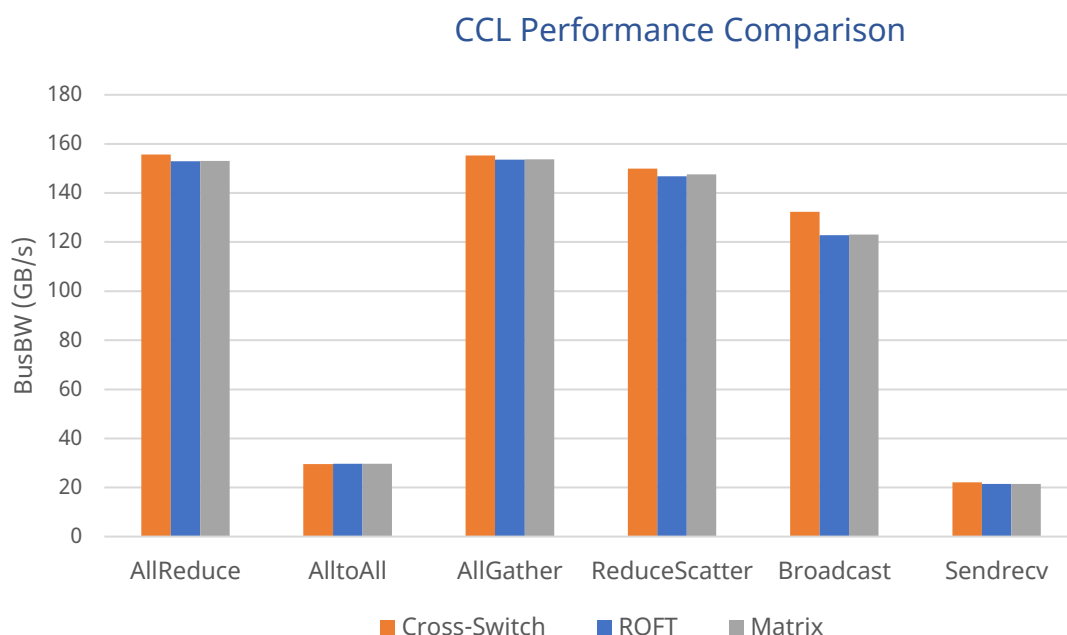


Figure 2 CCL Performance Comparison

The detailed test data is shown in the table below:

Table 2 CCL Performance Test Results

Network Type	BusBW (GB/s) ^[iv]					
	AllReduce	AlltoAll	AllGather	ReduceScatter	Broadcast	Sendrecv
Cross-Switch	155.66	29.58	155.28	149.94	132.28	22.17
AsterFabric ROFT	152.93	29.67	153.54	146.78	122.79	21.45
AsterFabric Matrix	153.01	29.69	153.73	147.58	123.01	21.44

After parameter tuning for the Matrix network, the performance of test items such as All-Reduce and All-to-All improved further. Matrix All-Reduce performance reached 162.78 GB/s and All-to-All performance reached 30.11 GB/s. The test data is as follows:

Table 3 Matrix Network Post-Tuning CCL Performance Test Results

Network Type	BusBW (GB/s)					
	AllReduce	AlltoAll	AllGather	ReduceScatter	Broadcast	Sendrecv
AsterFabric Matrix	162.78	30.11	160.25	155.25	125.82	30.33

3. In the model inference performance test (DeepSeek-V4-Flash model, 4 nodes with 32 GPUs, 2P2D), Matrix improved total throughput over ROFT by 10.5%, 24.6% and 21.7% at concurrency levels of 16, 32 and 64 respectively, and both Time To First Token (TTFT) and Time Per Output Token (TPOT) improved. See the figure below:

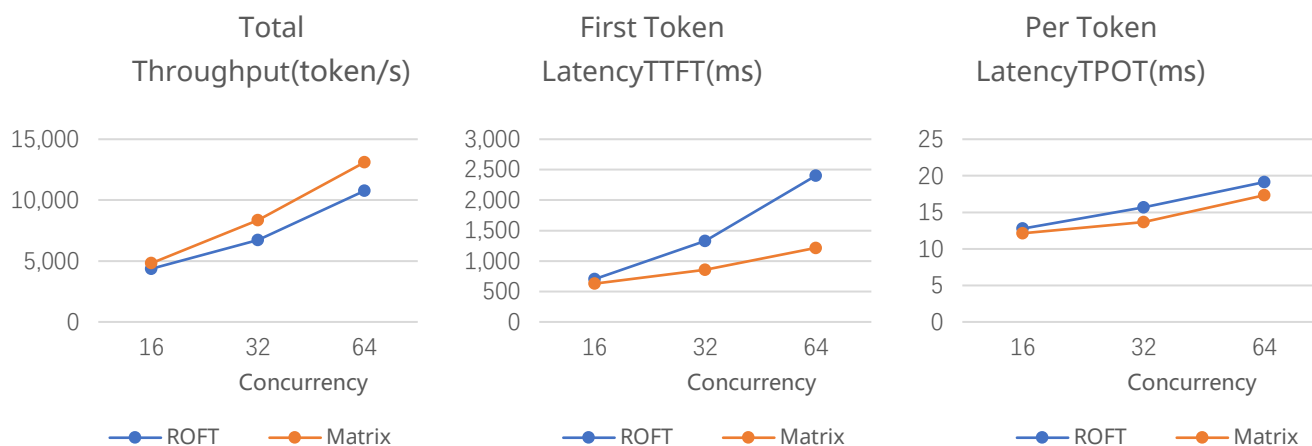


Figure 3 Model Inference Performance Comparison: TTFT and TPOT and Throughput

The detailed test data is as follows:

Table 4 Model Inference Performance Test Results

Concur ency	Total Throughput (token/s)		TTFT (ms)		TPOT (ms)	
	ROFT	Matrix	ROFT	Matrix	ROFT	Matrix
16	4,366.97	4,825.18	702.50	628.72	12.77	12.13
32	6,714.27	8,369.02	1,327.98	853.88	15.69	13.66
64	10,765.62	13,101.82	2,403.08	1,215.97	19.14	17.34

In summary, the test results show that both the AsterFabric ROFT and AsterFabric Matrix network architectures perform well in E2E forwarding and collective communication scenarios. In the 2P2D DeepSeek-V4-Flash model inference test, the Matrix architecture delivers higher total throughput and lower response latency.

Relying on Asterfusion's flexible network solution orchestration and end-to-end deployment capabilities, the network can be switched from ROFT to Matrix in a short time, with network deployment, service configuration and parameter tuning completed quickly, simplifying the overall delivery process. In the DeepSeek model inference test, the total throughput of the Matrix network at concurrency 64 reached 13,101.82 token/s, 21.7% higher than ROFT; Time To First Token (TTFT) was reduced by about 50% and Time Per Output Token (TPOT) by about 10%. In addition, based on a comparison of solution configurations, the network hardware cost and overall power consumption of the Matrix network can be reduced by about 30%; the actual benefits depend on the actual deployment configuration and calculation results.

Recommendations for solution selection are as follows:

- **Prioritize AsterFabric ROFT:** This architecture aligns well with current mainstream AI cluster

network design practices, and the related network design and engineering implementation experience is relatively mature. For customers who are unfamiliar with their workload models, lack the ability to tune GPU clusters and communication parameters, and are not yet in a position to do joint tuning, ROFT is recommended first, to reduce the complexity of implementation and tuning.

- **Prioritize evaluating AsterFabric Matrix:** For inference clusters that mainly adopt Prefill-Decode (PD) disaggregation and involve a large amount of cross-node KV Cache transfer, Matrix can be evaluated first. At the same GPU scale, this solution uses 33% fewer switches than ROFT, which can reduce hardware deployment costs. It should be noted that the Matrix network architecture increases implementation complexity to a certain extent and places higher demands on network planning, cable deployment and subsequent operations and maintenance. In addition, to fully realize its performance advantages, joint tuning that takes into account GPU cluster deployment, the communication framework and the characteristics of the inference workload is also required.

Overall, both the AsterFabric ROFT and AsterFabric Matrix architectures are viable networking options for AI computing networks. It is recommended to make the selection based on the actual service traffic model, performance goals and total cost of ownership (TCO).

Note:

[i] Total Throughput refers to the total number of output tokens produced jointly by all concurrent inference requests per unit time, commonly measured in token/s.

[ii] TTFT (Time To First Token) refers to the time elapsed from the moment a user sends a request until the model generates and returns the first token, commonly measured in ms.

[iii] TPOT (Time Per Output Token) refers to the average time required to generate each subsequent token when producing an output sequence, usually excluding the time to generate the first token, commonly measured in ms.

[iv] BusBW (Bus Bandwidth) is commonly used to measure the efficiency of a collective communication operation, i.e. the amount of data actually transferred by the algorithm relative to execution time, converted into an equivalent bandwidth based on the traffic volume of the communication algorithm.

2 Test Environment

2.1 Test Topology

2.1.1 Baseline Test Topology

In this topology, back-to-back direct NIC connection is used as the baseline test, with the NICs in 400G bond mode.

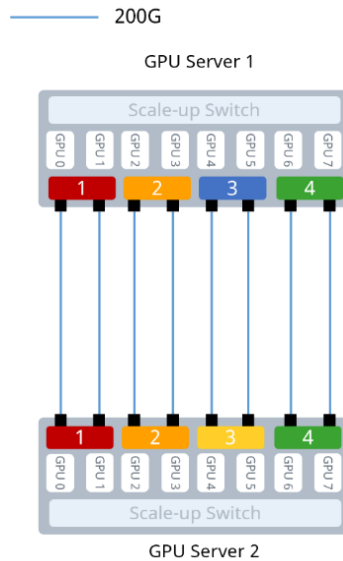


Figure 4 Baseline Test Topology

2.1.2 Cross-Switch Test Topology

This topology verifies whether the core performance and flow control behavior of a single switch under AI communication workloads meet design expectations, and ensures that the switch does not become a performance bottleneck for GPU communication. The NICs are in single-port 400G mode, and the result can serve as the reference baseline for the ROFT network.

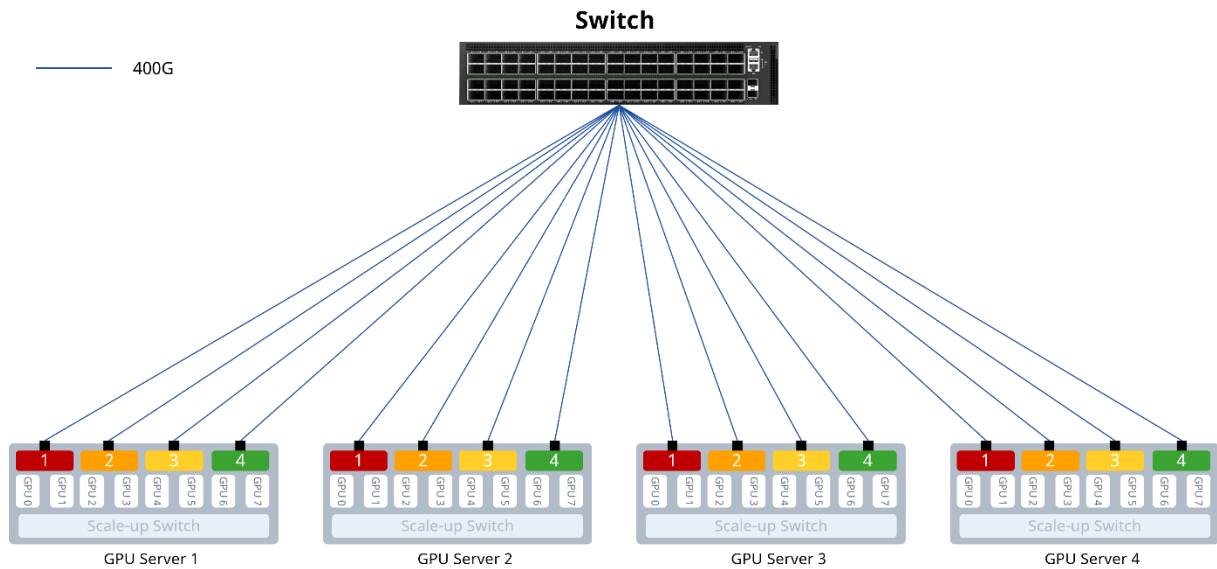


Figure 5 Cross-Switch Test Topology

2.1.3 AsterFabric ROFT Test Topology

This topology verifies GPU communication performance under the AsterFabric ROFT (Rail-Optimized Fat-Tree) architecture, with the NICs in single-port 400G mode.

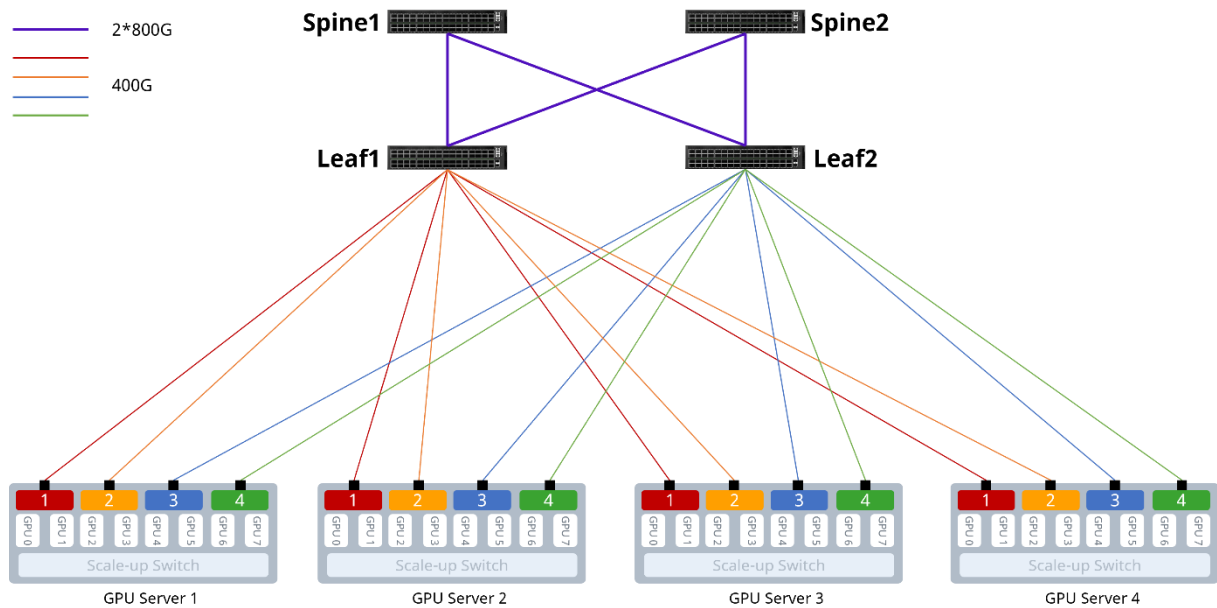


Figure 6 AsterFabric ROFT (Rail-Optimized) Test Topology

2.1.4 AsterFabric Matrix Test Topology

This topology verifies GPU communication performance under the AsterFabric Matrix (Orthogonal Dual-Plane) architecture, with the NICs in 400G bond mode.

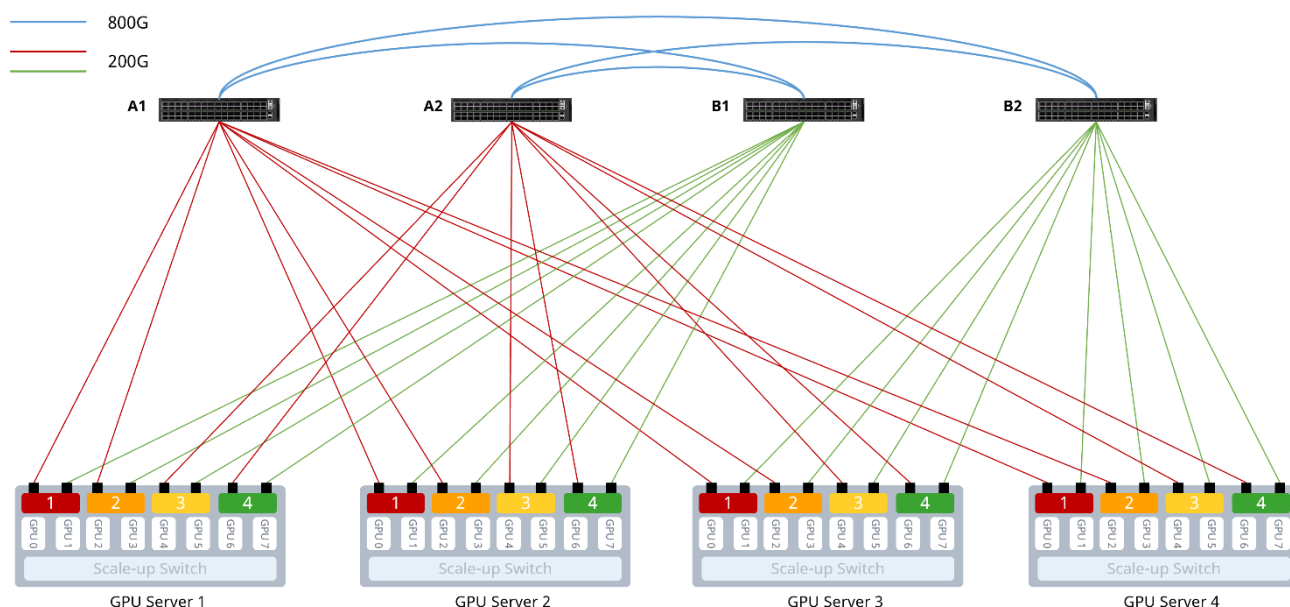


Figure 7 AsterFabric Matrix (Orthogonal Dual-Plane) Test Topology

This topology removes the Spine layer of the traditional Fat-Tree architecture and divides all Leaf switches evenly into two equally sized groups: Plane A (multi-rail plane, odd-numbered Leaf group) and Plane B (single-rail plane, even-numbered Leaf group). The two groups of switches are fully interconnected (that is, any switch in Group A has a direct link to every switch in Group B), while switches within the same plane are not connected to each other. The dual-port NIC of each GPU connects to both groups of switches: port 1 connects to a Plane A Leaf switch following the multi-rail rule, and port 2 connects to a Plane B Leaf switch following the single-rail rule.

2.2 Hardware and Software Environment

2.2.1 Switches

Table 5 Switch Information

Vendor	Model	Version	Qty.	Key Configuration Notes
Asterfusion	CX864E-N	R0409	4	CX864E-N data center switch, 64×800G/128×400G

2.2.2 Compute Nodes

There are 4 servers in total. Each server is equipped with 8 GPU cards, corresponding to 4 dual-port 200G NICs. The specific configuration is as follows:

Table 6 Compute Node Information

Category	Model/Configuration
CPU	INTEL(R) XEON(R) GOLD 6530
NIC	Broadcom BCM57608 2x200G PCIe Ethernet NIC
Operating System	Ubuntu 22.04.5 LTS / Linux 5.15.0-119-generic

NIC Driver	PWPKG Version: 235.2.97.0, L2DEV Version: 235.2.96.0
GPU Model	A domestic GPU ^[Note1]
GPU Compute Software Stack Version	
Collective Communication Library Version	
GPU Driver Version	
vBIOS Version	

2.2.3 Optical Modules and Cables

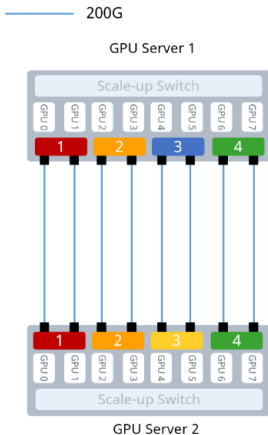
Table 7 Optical Module and Cable Information

Vendor	Type	Model	Qty.	Description
Asterfusion	800G optical module	800G-OSFP112-2VR4	24	800G-2VR4, OSFP112, multimode, dual-port MPO12/APC, 8-fiber, 30m/OM3, 50m/OM4
Asterfusion	400G optical module	400G-QSFP112-VR4	16	400G-VR4, QSFP112, multimode, MPO-12, 50m/OM4
Asterfusion	Optical fiber	MPO12/12S/APC-OM4	32	1*MPO12 to 1*MPO12, 12-fiber, APC connector, multimode fiber, OM4
Asterfusion	Optical fiber	MPO12/APC-2*MPO12/APC-OM4	16	1*MPO12/APC to 2*MPO12/APC, multimode fiber, OM4

[Note1] For further information, please contact an Asterfusion sales representative.

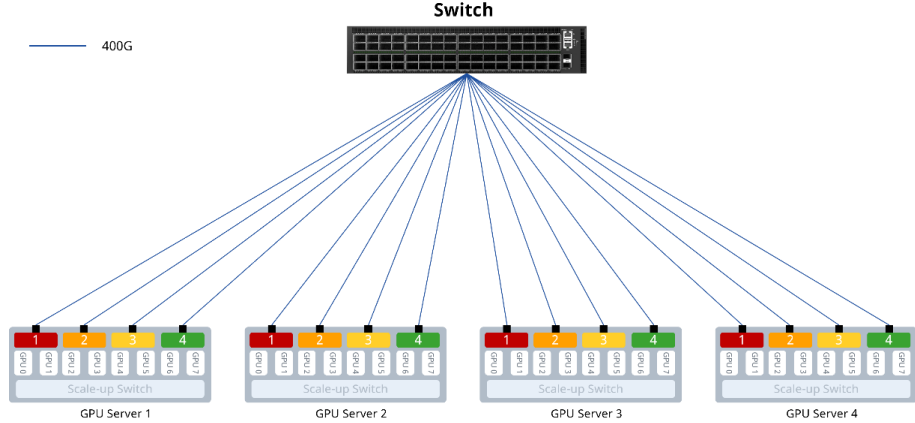
3 E2E Forwarding Performance Test

3.1 NIC Back-to-Back Forwarding Bandwidth Test

Test Case	Server NIC Back-to-Back Forwarding Bandwidth Test
Test Objective	Two servers are connected directly NIC to NIC, and the bandwidth performance of any pair of NICs is tested.
Test Topology	
Test Procedure	1. Select any two servers and connect them back to back directly. 2. Use the ib_write_bw tool to test the bandwidth of any directly connected pair of NICs; see Expected Result 1.
Expected Result	1. Record the bandwidth test result and verify whether it matches the theoretical value.
Test Result	1. The back-to-back bandwidth is approximately 390.62 Gb/s. <pre> RDMA_Write BW Test Dual-port : OFF Device : roce2 Number of qps : 2 Transport type : IB Connection type : RC Using SRQ : OFF PCIe relax order: ON ibv_wr* API : OFF TX depth : 128 CQ Moderation : 1 Mtu : 4096[B] Link type : Ethernet GID index : 3 Max inline data : 0[B] rdma_cm QPs : OFF Data ex. method : Ethernet ----- local address: LID 0000 QPN 0x0401 PSN 0x38c0ce RKey 0x4000001 VAddr 0x007fafb2420000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:03:06 local address: LID 0000 QPN 0x0402 PSN 0x1295a0 RKey 0x4000001 VAddr 0x007fafb2430000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:03:06 remote address: LID 0000 QPN 0x0401 PSN 0x7757fe RKey 0x4000001 VAddr 0x007fda51020000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:03:05 remote address: LID 0000 QPN 0x0402 PSN 0xce1510 RKey 0x4000001 VAddr 0x007fda51030000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:03:05 ----- #bytes #iterations BW peak[Gb/sec] BW average[Gb/sec] MsgRate[Mpps] 65536 744960 0.00 390.54 0.744888 65536 745100 0.00 390.61 0.745033 65536 745106 0.00 390.62 0.745041 </pre>
Test	Meets expectations

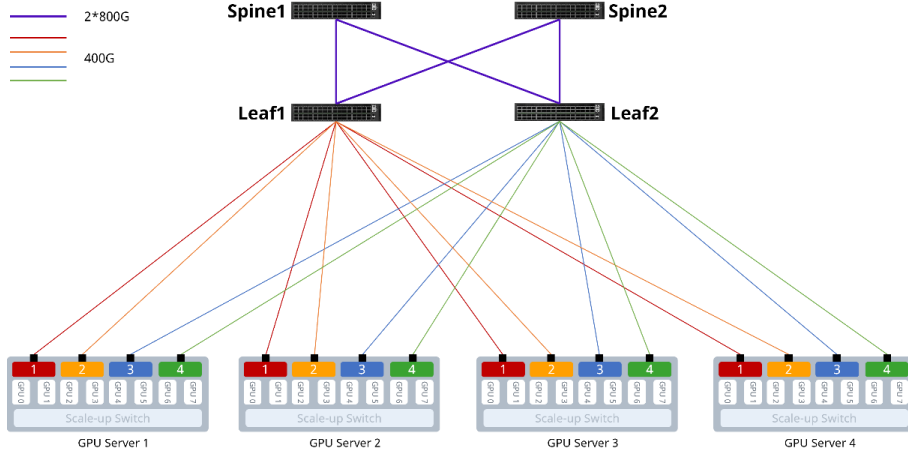
Conclusion	
------------	--

3.2 Cross-Switch Forwarding Bandwidth Test

Test Case	Cross-Switch Forwarding Bandwidth Test
Test Objective	Four servers are connected to a single switch, and the bandwidth performance of any pair of NICs is tested.
Test Topology	
Test Procedure	1. Connect four servers to a single switch. 2. Use the ib_write_bw tool to test the bandwidth of any directly connected pair of NICs; see Expected Result 1.
Expected Result	1. Record the bandwidth test result and verify whether it matches the theoretical value.
Test Result	1. The bandwidth forwarded across a single switch is approximately 392.05 Gb/s. <pre> RDMA_Write BW Test Dual-port : OFF Device : bnxt_re0 Number of qps : 2 Transport type : IB Connection type : RC Using SRQ : OFF PCIe relax order: ON ibv_wr* API : OFF TX depth : 128 CQ Moderation : 1 Mtu : 4096[B] Link type : Ethernet GID index : 3 Max inline data : 0[B] rdma_cm QPs : OFF Data ex. method : Ethernet ----- local address: LID 0000 QPN 0x11be PSN 0xc84e57 RKey 0x4021169 VAddr 0x007fc31b820000 GID: 00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:13 local address: LID 0000 QPN 0x0514 PSN 0x1d622c RKey 0x4021169 VAddr 0x007fc31b830000 GID: 00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:13 remote address: LID 0000 QPN 0x069b PSN 0x77388c RKey 0x4050b41 VAddr 0x007f502bc20000 GID: 00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:09 remote address: LID 0000 QPN 0x07ce PSN 0xdeb8bd RKey 0x4050b41 VAddr 0x007f502bc30000 GID: 00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:09 ----- #bytes #iterations BW peak[Gb/sec] BW average[Gb/sec] MsgRate[Mpps] 65536 747804 0.00 392.02 0.747728 65536 747820 0.00 392.05 0.747779 65536 747801 0.00 392.05 0.747783 65536 747798 0.00 392.05 0.747775 RESULT: bnxt_re0 GPU0 PASS BW#1=392.02 BW#2=392.05 BW#3=392.05 Gb/s </pre>

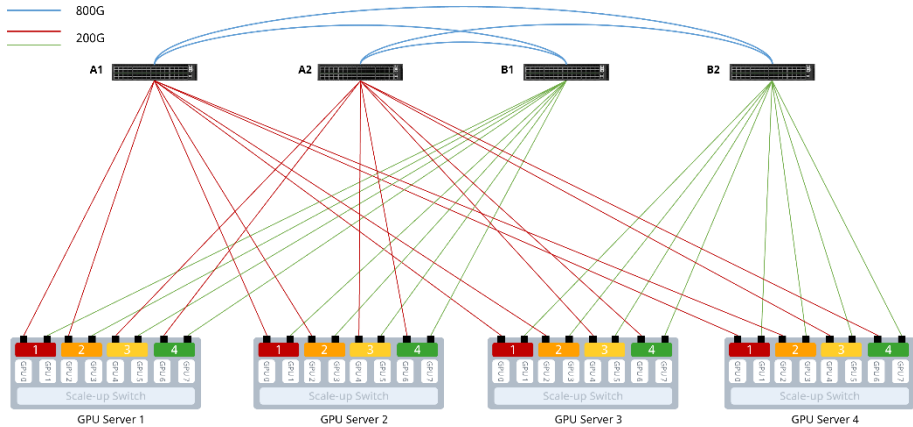
Test Conclusion	Meets expectations
------------------------	--------------------

3.3 AsterFabric ROFT Forwarding Bandwidth Test

Test Case	AsterFabric ROFT Forwarding Bandwidth Test
Test Objective	Test the bandwidth performance under the AsterFabric ROFT architecture.
Test Topology	
Test Procedure	- Build the network according to the AsterFabric ROFT architecture. - Use the <code>ib_write_bw</code> tool to test the bandwidth of any directly connected pair of NICs; see Expected Result 1.
Expected Result	Record the bandwidth test result and verify whether it is comparable to the reference baseline (cross-switch) performance.
Test Result	The bandwidth under the ROFT network is approximately 392.06 Gb/s. <pre> RDMA_Write BW Test Dual-port : OFF Device : bnxt_re0 Number of qps : 2 Transport type : IB Connection type : RC Using SRQ : OFF PCIe relax order: ON ibv_wr* API : OFF TX depth : 128 CQ Moderation : 1 MTU : 4096[B] Link type : Ethernet GID index : 3 Max inline data : 0[B] rdma_cm QPs : OFF Data ex. method : Ethernet ----- local address: LID 0000 QPN 0x0401 PSN 0x8611ee RKey 0x4000003 VAddr 0x007f1b29820000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:13 local address: LID 0000 QPN 0x0402 PSN 0x60e040 RKey 0x4000003 VAddr 0x007f1b29830000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:13 remote address: LID 0000 QPN 0x0401 PSN 0x7a1763 RKey 0x4000003 VAddr 0x007f5460820000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:09 remote address: LID 0000 QPN 0x0402 PSN 0xa5111 RKey 0x4000003 VAddr 0x007f5460830000 GID: 00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:09 ----- #bytes #iterations BW peak[Gb/sec] BW average[Gb/sec] MsgRate[Mpps] 65536 747779 0.00 392.02 0.747721 65536 747820 0.00 392.04 0.747755 65536 747822 0.00 392.06 0.747787 65536 747814 0.00 392.06 0.747787 RESULT: bnxt_re0 GPU0 PASS BW#1=392.02 BW#2=392.04 BW#3=392.06 Gb/s </pre>

Test Conclusion	Meets expectations
------------------------	--------------------

3.4 AsterFabric Matrix Forwarding Bandwidth Test

Test Case	AsterFabric Matrix Forwarding Bandwidth Test
Test Objective	Test the bandwidth performance under the AsterFabric Matrix architecture.
Test Topology	
Test Procedure	- Build the network according to the AsterFabric Matrix architecture. - Use the <code>ib_write_bw</code> tool to test the bandwidth of any directly connected pair of NICs; see Expected Result 1.
Expected Result	Record the bandwidth test result and verify whether it is comparable to the reference baseline (NIC back-to-back) performance.
Test Result	The bandwidth under the Matrix network is approximately 390.63 Gb/s. <pre> RDMA_Write BW Test Dual-port : OFF Device : roce0 Number of qps : 2 Transport type : IB Connection type: RC Using SRQ : OFF PCIe relax order: ON ibv_wr* API : OFF TX depth : 128 CQ Moderation : 1 Mtu : 4096[B] Link type : Ethernet GID index : 3 Max inline data: 0[B] rdma_cm QPs : OFF Data ex. method : Ethernet ----- local address: LID 0000 QPN 0x0508 PSN 0x881b48 RKey 0x400637c VAddr 0x007fb354020000 GID: 00:00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:13 local address: LID 0000 QPN 0x0497 PSN 0x9db792 RKey 0x400637c VAddr 0x007fb354030000 GID: 00:00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:13 remote address: LID 0000 QPN 0x04e2 PSN 0xa67c46 RKey 0x400a82f VAddr 0x007fa28f820000 GID: 00:00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:09 remote address: LID 0000 QPN 0x04f2 PSN 0x4faaef RKey 0x400a82f VAddr 0x007fa28f830000 GID: 00:00:00:00:00:00:00:00:00:00:00:00:255:255:10:01:01:09 ----- #bytes #iterations BW peak[Gb/sec] BW average[Gb/sec] MsgRate[Mpps] 65536 744997 0.00 390.56 0.744939 65536 745138 0.00 390.63 0.745072 65536 745131 0.00 390.63 0.745071 65536 745138 0.00 390.65 0.745100 RESULT: roce0 GPU0 PASS BW#1=390.56 BW#2=390.63 BW#3=390.63 Gb/s </pre>

Test Conclusion	Meets expectations
------------------------	--------------------

4 Collective Communication Performance Test

4.1 Cross-Switch CCL Test

4.1.1 Cross-Switch All-Reduce Test

Test Case	Cross-Switch All-Reduce Test
Test Objective	Test the performance of the All-Reduce communication model in a single-switch network.
Test Topology	See 2.1.2
Test Procedure	1. Build the network according to the single-switch architecture. 2. Use the CCL script to test All-Reduce performance on 4 nodes with 32 GPUs; see Expected Result 1.
Expected Result	1. Record the out-of-place busbw and verify whether it matches the theoretical value.
Test Result	1. In the single-switch network, the bus bandwidth of All-Reduce at 8G size is 155.66 GB/s.

	<pre># Using devices # Rank 0 Pid 443721 on CX7Group1-host-061 device 0 [0x2b] # Rank 1 Pid 443722 on CX7Group1-host-061 device 1 [0x3a] # Rank 2 Pid 443723 on CX7Group1-host-061 device 2 [0x4d] # Rank 3 Pid 443726 on CX7Group1-host-061 device 3 [0x5c] # Rank 4 Pid 443727 on CX7Group1-host-061 device 4 [0xaa] # Rank 5 Pid 443728 on CX7Group1-host-061 device 5 [0xba] # Rank 6 Pid 443729 on CX7Group1-host-061 device 6 [0xca] # Rank 7 Pid 443730 on CX7Group1-host-061 device 7 [0xda] # Rank 8 Pid 1799340 on CX7Group1-host-062 device 0 [0x2b] # Rank 9 Pid 1799341 on CX7Group1-host-062 device 1 [0x3a] # Rank 10 Pid 1799342 on CX7Group1-host-062 device 2 [0x4d] # Rank 11 Pid 1799343 on CX7Group1-host-062 device 3 [0x5c] # Rank 12 Pid 1799344 on CX7Group1-host-062 device 4 [0xaa] # Rank 13 Pid 1799345 on CX7Group1-host-062 device 5 [0xba] # Rank 14 Pid 1799346 on CX7Group1-host-062 device 6 [0xca] # Rank 15 Pid 1799347 on CX7Group1-host-062 device 7 [0xda] # Rank 16 Pid 3076717 on CX7Group1-host-063 device 0 [0x2b] # Rank 17 Pid 3076719 on CX7Group1-host-063 device 1 [0x3a] # Rank 18 Pid 3076721 on CX7Group1-host-063 device 2 [0x4d] # Rank 19 Pid 3076723 on CX7Group1-host-063 device 3 [0x5c] # Rank 20 Pid 3076725 on CX7Group1-host-063 device 4 [0xaa] # Rank 21 Pid 3076727 on CX7Group1-host-063 device 5 [0xba] # Rank 22 Pid 3076729 on CX7Group1-host-063 device 6 [0xca] # Rank 23 Pid 3076733 on CX7Group1-host-063 device 7 [0xda] # Rank 24 Pid 182335 on CX7Group1-host-064 device 0 [0x2b] # Rank 25 Pid 182336 on CX7Group1-host-064 device 1 [0x3a] # Rank 26 Pid 182337 on CX7Group1-host-064 device 2 [0x4d] # Rank 27 Pid 182338 on CX7Group1-host-064 device 3 [0x5c] # Rank 28 Pid 182339 on CX7Group1-host-064 device 4 [0xaa] # Rank 29 Pid 182340 on CX7Group1-host-064 device 5 [0xba] # Rank 30 Pid 182341 on CX7Group1-host-064 device 6 [0xca] # Rank 31 Pid 182342 on CX7Group1-host-064 device 7 [0xda] # # # 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) 1024 256 float sum -1 61.89 0.02 0.03 0 58.54 0.02 0.03 0 2048 512 float sum -1 62.23 0.03 0.06 0 62.60 0.03 0.06 0 4096 1024 float sum -1 68.48 0.06 0.12 0 67.50 0.06 0.12 0 8192 2048 float sum -1 71.14 0.12 0.22 0 68.80 0.12 0.23 0 16384 4096 float sum -1 70.74 0.23 0.45 0 68.61 0.24 0.46 0 32768 8192 float sum -1 73.65 0.44 0.86 0 71.65 0.46 0.89 0 65536 16384 float sum -1 79.11 0.83 1.61 0 74.08 0.88 1.71 0 131072 32768 float sum -1 179.59 0.73 1.41 0 177.13 0.74 1.43 0 262144 65536 float sum -1 182.52 1.44 2.78 0 183.10 1.43 2.77 0 524288 131072 float sum -1 194.40 2.70 5.23 0 200.51 2.61 5.07 0 1048576 262144 float sum -1 230.83 4.54 8.80 0 221.67 4.73 9.16 0 2097152 524288 float sum -1 246.90 8.49 16.46 0 245.68 8.54 16.54 0 4194304 1048576 float sum -1 296.88 14.13 27.37 0 295.72 14.18 27.48 0 8388608 2097152 float sum -1 394.21 21.28 41.23 0 394.05 21.29 41.25 0 16777216 4194304 float sum -1 632.35 26.53 51.40 0 615.51 27.26 52.81 0 33554432 8388608 float sum -1 868.97 38.61 74.82 0 867.07 38.70 74.98 0 67108864 16777216 float sum -1 1425.65 47.07 91.20 0 1406.95 47.70 92.42 0 134217728 33554432 float sum -1 2836.47 47.32 91.68 0 2832.91 47.38 91.79 0 268435456 67108864 float sum -1 5121.12 52.42 101.56 0 5145.78 52.17 101.07 0 536870912 134217728 float sum -1 8239.96 65.15 126.24 0 8197.30 65.49 126.89 0 1073741824 268435456 float sum -1 15421.49 69.63 134.90 0 15353.58 69.93 135.50 0 2147483648 536870912 float sum -1 28547.88 75.22 145.75 0 28585.56 75.12 145.55 0 4294967296 1073741824 float sum -1 53381.15 80.46 155.89 0 53316.64 80.56 156.08 0 8589934592 2147483648 float sum -1 106915.94 80.34 155.66 0 106838.48 80.40 155.78 0 # Out of bounds values : 0 OK # Avg bus bandwidth : 51.5795</pre>
Test Conclusion	Meets expectations
Test Notes	Pre-tuning parameters are used; the bus bandwidth at an 8G message size is taken as the test result.

4.1.2 Cross-Switch All-to-All Test

Test Case	Cross-Switch All-to-All Test
Test Objective	Test the performance of the All-to-All communication model in a single-switch network.
Test Topology	See 2.1.2
Test Procedure	1. Build the network according to the single-switch architecture. 2. Use the CCL script to test All-to-All performance on 4 nodes with 32 GPUs; see Expected Result 1.
Expected Result	1. Record the out-of-place busbw and verify whether it matches the theoretical value.
Test Result	1. In the single-switch network, the bus bandwidth of All-to-All at 8G size is 29.58 GB/s.

	<pre> # Using devices # Rank 0 Pid 532471 on CX7Group1-host-061 device 0 [0x2b] # Rank 1 Pid 532472 on CX7Group1-host-061 device 1 [0x3a] # Rank 2 Pid 532473 on CX7Group1-host-061 device 2 [0x4d] # Rank 3 Pid 532474 on CX7Group1-host-061 device 3 [0x5c] # Rank 4 Pid 532475 on CX7Group1-host-061 device 4 [0xaa] # Rank 5 Pid 532476 on CX7Group1-host-061 device 5 [0xba] # Rank 6 Pid 532477 on CX7Group1-host-061 device 6 [0xca] # Rank 7 Pid 532478 on CX7Group1-host-061 device 7 [0xda] # Rank 8 Pid 1908912 on CX7Group1-host-062 device 0 [0x2b] # Rank 9 Pid 1908913 on CX7Group1-host-062 device 1 [0x3a] # Rank 10 Pid 1908914 on CX7Group1-host-062 device 2 [0x4d] # Rank 11 Pid 1908915 on CX7Group1-host-062 device 3 [0x5c] # Rank 12 Pid 1908916 on CX7Group1-host-062 device 4 [0xaa] # Rank 13 Pid 1908920 on CX7Group1-host-062 device 5 [0xba] # Rank 14 Pid 1908922 on CX7Group1-host-062 device 6 [0xca] # Rank 15 Pid 1908925 on CX7Group1-host-062 device 7 [0xda] # Rank 16 Pid 3182622 on CX7Group1-host-063 device 0 [0x2b] # Rank 17 Pid 3182623 on CX7Group1-host-063 device 1 [0x3a] # Rank 18 Pid 3182624 on CX7Group1-host-063 device 2 [0x4d] # Rank 19 Pid 3182625 on CX7Group1-host-063 device 3 [0x5c] # Rank 20 Pid 3182626 on CX7Group1-host-063 device 4 [0xaa] # Rank 21 Pid 3182627 on CX7Group1-host-063 device 5 [0xba] # Rank 22 Pid 3182628 on CX7Group1-host-063 device 6 [0xca] # Rank 23 Pid 3182629 on CX7Group1-host-063 device 7 [0xda] # Rank 24 Pid 270764 on CX7Group1-host-064 device 0 [0x2b] # Rank 25 Pid 270765 on CX7Group1-host-064 device 1 [0x3a] # Rank 26 Pid 270766 on CX7Group1-host-064 device 2 [0x4d] # Rank 27 Pid 270767 on CX7Group1-host-064 device 3 [0x5c] # Rank 28 Pid 270768 on CX7Group1-host-064 device 4 [0xaa] # Rank 29 Pid 270769 on CX7Group1-host-064 device 5 [0xba] # Rank 30 Pid 270770 on CX7Group1-host-064 device 6 [0xca] # Rank 31 Pid 270771 on CX7Group1-host-064 device 7 [0xda] # # # size count type redop root time algbw busbw #wrong time algbw busbw #wrong # (B) (elements) # 1024 8 float -1 160.40 0.01 0.01 0 145.54 0.01 0.01 N/A # 2048 16 float -1 146.31 0.01 0.01 0 148.21 0.01 0.01 N/A # 4096 32 float -1 145.81 0.03 0.03 0 145.50 0.03 0.03 N/A # 8192 64 float -1 146.48 0.06 0.05 0 146.64 0.06 0.05 N/A # 16384 128 float -1 146.59 0.11 0.11 0 146.43 0.11 0.11 N/A # 32768 256 float -1 148.21 0.22 0.21 0 146.54 0.22 0.22 N/A # 65536 512 float -1 149.96 0.44 0.42 0 149.98 0.44 0.42 N/A # 131072 1024 float -1 152.79 0.86 0.83 0 154.29 0.85 0.82 N/A # 262144 2048 float -1 155.36 1.69 1.63 0 175.02 1.49 1.44 N/A # 524288 4096 float -1 175.26 2.99 2.90 0 166.86 3.14 3.04 N/A # 1048576 8192 float -1 198.19 5.29 5.13 0 195.22 5.37 5.20 N/A # 2097152 16384 float -1 256.97 8.16 7.91 0 253.60 8.27 8.01 N/A # 4194304 32768 float -1 340.51 12.32 11.93 0 338.60 12.39 12.00 N/A # 8388608 65536 float -1 485.54 17.28 16.74 0 485.90 17.26 16.72 N/A # 16777216 131072 float -1 866.61 19.36 18.75 0 863.01 19.44 18.83 N/A # 33554432 262144 float -1 1406.30 23.86 23.11 0 1405.22 23.88 23.13 N/A # 67108864 524288 float -1 2675.00 25.09 24.30 0 2690.37 24.94 24.16 N/A # 134217728 1048576 float -1 4628.40 29.00 28.09 0 4670.64 28.74 27.84 N/A # 268435456 2097152 float -1 9299.82 28.86 27.96 0 9026.06 29.74 28.81 N/A # 536870912 4194304 float -1 17862.48 30.06 29.12 0 17827.07 30.12 29.17 N/A # 1073741824 8388608 float -1 35477.66 30.27 29.32 0 35859.61 29.94 29.01 N/A # 2147483648 16777216 float -1 70879.41 30.30 29.35 0 70972.07 30.26 29.31 N/A # 4294967296 33554432 float -1 141411.99 30.37 29.42 0 141327.00 30.39 29.44 N/A # 8589934592 67108864 float -1 281341.73 30.53 29.58 0 281464.87 30.52 29.56 N/A # Out of bounds values : 0 OK # Avg bus bandwidth : 13.2147 </pre>
Test Conclusion	Meets expectations
Test Notes	Pre-tuning parameters are used; the bus bandwidth at an 8G message size is taken as the test result.

4.2 AsterFabric ROFT CCL Test

4.2.1 AsterFabric ROFT All-Reduce Test

Test Case	AsterFabric ROFT All-Reduce Test
Test Objective	Test the performance of the All-Reduce communication model under the AsterFabric ROFT architecture.
Test Topology	See 2.1.3
Test Procedure	- Build the network according to the AsterFabric ROFT architecture. - Use the CCL script to test All-Reduce performance on 4 nodes with 32 GPUs; see Expected Result 1.

Expected Result	1. Record the out-of-place busbw and verify whether it is comparable to the reference baseline performance.
Test Result	1. In the ROFT network, the bus bandwidth of All-Reduce at 8G size is 152.93 GB/s. <pre># Using devices # Rank 0 Pid 373601 on CX7Group1-host-061 device 0 [0x2b] # Rank 1 Pid 373602 on CX7Group1-host-061 device 1 [0x3a] # Rank 2 Pid 373603 on CX7Group1-host-061 device 2 [0x4d] # Rank 3 Pid 373604 on CX7Group1-host-061 device 3 [0x5c] # Rank 4 Pid 373605 on CX7Group1-host-061 device 4 [0xaa] # Rank 5 Pid 373606 on CX7Group1-host-061 device 5 [0xba] # Rank 6 Pid 373607 on CX7Group1-host-061 device 6 [0xca] # Rank 7 Pid 373608 on CX7Group1-host-061 device 7 [0xda] # Rank 8 Pid 858751 on CX7Group1-host-062 device 0 [0x2b] # Rank 9 Pid 858752 on CX7Group1-host-062 device 1 [0x3a] # Rank 10 Pid 858753 on CX7Group1-host-062 device 2 [0x4d] # Rank 11 Pid 858754 on CX7Group1-host-062 device 3 [0x5c] # Rank 12 Pid 858755 on CX7Group1-host-062 device 4 [0xaa] # Rank 13 Pid 858756 on CX7Group1-host-062 device 5 [0xba] # Rank 14 Pid 858757 on CX7Group1-host-062 device 6 [0xca] # Rank 15 Pid 858758 on CX7Group1-host-062 device 7 [0xda] # Rank 16 Pid 408630 on CX7Group1-host-063 device 0 [0x2b] # Rank 17 Pid 408631 on CX7Group1-host-063 device 1 [0x3a] # Rank 18 Pid 408632 on CX7Group1-host-063 device 2 [0x4d] # Rank 19 Pid 408633 on CX7Group1-host-063 device 3 [0x5c] # Rank 20 Pid 408634 on CX7Group1-host-063 device 4 [0xaa] # Rank 21 Pid 408635 on CX7Group1-host-063 device 5 [0xba] # Rank 22 Pid 408636 on CX7Group1-host-063 device 6 [0xca] # Rank 23 Pid 408637 on CX7Group1-host-063 device 7 [0xda] # Rank 24 Pid 379804 on CX7Group1-host-064 device 0 [0x2b] # Rank 25 Pid 379805 on CX7Group1-host-064 device 1 [0x3a] # Rank 26 Pid 379806 on CX7Group1-host-064 device 2 [0x4d] # Rank 27 Pid 379807 on CX7Group1-host-064 device 3 [0x5c] # Rank 28 Pid 379808 on CX7Group1-host-064 device 4 [0xaa] # Rank 29 Pid 379809 on CX7Group1-host-064 device 5 [0xba] # Rank 30 Pid 379810 on CX7Group1-host-064 device 6 [0xca] # Rank 31 Pid 379811 on CX7Group1-host-064 device 7 [0xda] # # # 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) # # 1024 256 float sum -1 63.20 0.02 0.03 0 59.97 0.02 0.03 0 # 2048 512 float sum -1 63.58 0.03 0.06 0 63.76 0.03 0.06 0 # 4096 1024 float sum -1 69.09 0.06 0.11 0 68.14 0.06 0.12 0 # 8192 2048 float sum -1 70.86 0.12 0.22 0 68.73 0.12 0.23 0 # 16384 4096 float sum -1 70.81 0.23 0.45 0 69.66 0.24 0.46 0 # 32768 8192 float sum -1 73.78 0.44 0.86 0 71.33 0.46 0.89 0 # 65536 16384 float sum -1 79.76 0.82 1.59 0 75.58 0.87 1.68 0 # 131072 32768 float sum -1 178.60 0.73 1.42 0 179.86 0.73 1.41 0 # 262144 65536 float sum -1 183.12 1.43 2.77 0 184.94 1.42 2.75 0 # 524288 131072 float sum -1 195.34 2.68 5.20 0 194.48 2.70 5.23 0 # 1048576 262144 float sum -1 235.17 4.46 8.64 0 223.82 4.68 9.08 0 # 2097152 524288 float sum -1 247.42 8.48 16.42 0 245.38 8.55 16.56 0 # 4194304 1048576 float sum -1 296.72 14.14 27.39 0 295.57 14.19 27.49 0 # 8388608 2097152 float sum -1 392.36 21.38 41.42 0 393.15 21.34 41.34 0 # 16777216 4194304 float sum -1 622.56 26.95 52.21 0 620.61 27.03 52.38 0 # 33554432 8388608 float sum -1 863.70 38.85 75.26 0 861.44 38.95 75.47 0 # 67108864 16777216 float sum -1 1410.13 47.59 92.21 0 1406.72 47.71 92.43 0 # 134217728 33554432 float sum -1 3071.90 34.66 67.16 0 3073.35 34.65 67.14 0 # 268435456 67108864 float sum -1 6441.99 41.67 80.73 0 6448.01 41.63 80.66 0 # 536870912 134217728 float sum -1 9088.65 59.60 115.47 0 9087.50 59.60 115.48 0 # 1073741824 268435456 float sum -1 16388.17 65.52 126.94 0 16400.59 65.47 126.85 0 # 2147483648 536870912 float sum -1 29443.26 72.94 141.31 0 29593.28 72.57 140.60 0 # 4294967296 1073741824 float sum -1 54369.50 79.00 153.05 0 54351.82 79.02 153.10 0 # 8589934592 2147483648 float sum -1 108828.90 78.93 152.93 0 108984.86 78.82 152.71 0 # # Out of bounds values : 0 OK # Avg bus bandwidth : 48.5006</pre>
Test Conclusion	Meets expectations
Test Notes	Pre-tuning parameters are used; the bus bandwidth at an 8G message size is taken as the test result.

4.2.2 AsterFabric ROFT All-to-All Test

Test Case	AsterFabric ROFT All-to-All Test
Test Objective	Test the performance of the All-to-All communication model under the AsterFabric ROFT architecture.
Test Topology	See 2.1.3
Test Procedure	1. Build the network according to the AsterFabric ROFT architecture. 2. Use the CCL script to test All-to-All performance on 4 nodes with 32 GPUs; see Expected Result 1.

Expected Result	1. Record the out-of-place busbw and verify whether it is comparable to the reference baseline performance.
Test Result	1. In the ROFT network, the bus bandwidth of All-to-All at 8G size is 29.67 GB/s. <pre> # Using devices # Rank 0 Pid 379341 on CX7Group1-host-061 device 0 [0x2b] # Rank 1 Pid 379342 on CX7Group1-host-061 device 1 [0x3a] # Rank 2 Pid 379343 on CX7Group1-host-061 device 2 [0x4d] # Rank 3 Pid 379344 on CX7Group1-host-061 device 3 [0x5c] # Rank 4 Pid 379345 on CX7Group1-host-061 device 4 [0xaa] # Rank 5 Pid 379346 on CX7Group1-host-061 device 5 [0xba] # Rank 6 Pid 379347 on CX7Group1-host-061 device 6 [0xca] # Rank 7 Pid 379348 on CX7Group1-host-061 device 7 [0xda] # Rank 8 Pid 867735 on CX7Group1-host-062 device 0 [0x2b] # Rank 9 Pid 867737 on CX7Group1-host-062 device 1 [0x3a] # Rank 10 Pid 867738 on CX7Group1-host-062 device 2 [0x4d] # Rank 11 Pid 867739 on CX7Group1-host-062 device 3 [0x5c] # Rank 12 Pid 867742 on CX7Group1-host-062 device 4 [0xaa] # Rank 13 Pid 867743 on CX7Group1-host-062 device 5 [0xba] # Rank 14 Pid 867744 on CX7Group1-host-062 device 6 [0xca] # Rank 15 Pid 867751 on CX7Group1-host-062 device 7 [0xda] # Rank 16 Pid 414472 on CX7Group1-host-063 device 0 [0x2b] # Rank 17 Pid 414473 on CX7Group1-host-063 device 1 [0x3a] # Rank 18 Pid 414474 on CX7Group1-host-063 device 2 [0x4d] # Rank 19 Pid 414475 on CX7Group1-host-063 device 3 [0x5c] # Rank 20 Pid 414476 on CX7Group1-host-063 device 4 [0xaa] # Rank 21 Pid 414477 on CX7Group1-host-063 device 5 [0xba] # Rank 22 Pid 414478 on CX7Group1-host-063 device 6 [0xca] # Rank 23 Pid 414479 on CX7Group1-host-063 device 7 [0xda] # Rank 24 Pid 385548 on CX7Group1-host-064 device 0 [0x2b] # Rank 25 Pid 385549 on CX7Group1-host-064 device 1 [0x3a] # Rank 26 Pid 385550 on CX7Group1-host-064 device 2 [0x4d] # Rank 27 Pid 385551 on CX7Group1-host-064 device 3 [0x5c] # Rank 28 Pid 385552 on CX7Group1-host-064 device 4 [0xaa] # Rank 29 Pid 385553 on CX7Group1-host-064 device 5 [0xba] # Rank 30 Pid 385554 on CX7Group1-host-064 device 6 [0xca] # Rank 31 Pid 385555 on CX7Group1-host-064 device 7 [0xda] # # # 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) # 1024 8 float -1 155.93 0.01 0.01 0 150.38 0.01 0.01 N/A 2048 16 float -1 147.11 0.01 0.01 0 153.66 0.01 0.01 N/A 4096 32 float -1 146.58 0.03 0.03 0 145.46 0.03 0.03 N/A 8192 64 float -1 147.89 0.06 0.05 0 146.19 0.06 0.05 N/A 16384 128 float -1 210.73 0.08 0.08 0 148.52 0.11 0.11 N/A 32768 256 float -1 147.03 0.22 0.22 0 147.96 0.22 0.21 N/A 65536 512 float -1 151.11 0.43 0.42 0 150.59 0.44 0.42 N/A 131072 1024 float -1 149.98 0.87 0.85 0 147.09 0.89 0.86 N/A 262144 2048 float -1 155.04 1.69 1.64 0 156.12 1.68 1.63 N/A 524288 4096 float -1 167.07 3.14 3.04 0 165.84 3.16 3.06 N/A 1048576 8192 float -1 194.43 5.39 5.22 0 200.89 5.22 5.06 N/A 2097152 16384 float -1 250.83 8.36 8.10 0 250.33 8.38 8.12 N/A 4194304 32768 float -1 337.33 12.43 12.05 0 338.35 12.48 12.01 N/A 8388608 65536 float -1 479.14 17.51 16.96 0 472.95 17.74 17.18 N/A 16777216 131072 float -1 804.40 20.86 20.28 0 807.01 20.79 20.14 N/A 33554432 262144 float -1 1305.35 24.22 23.46 0 1409.42 23.81 23.06 N/A 67108864 524288 float -1 2540.75 26.41 25.59 0 2529.37 26.53 25.70 N/A 134217728 1048576 float -1 4653.40 28.84 27.94 0 4671.09 28.73 27.84 N/A 268435456 2097152 float -1 9186.55 29.22 28.31 0 9284.86 29.16 28.25 N/A 536870912 4194304 float -1 17943.79 29.92 28.98 0 17852.34 30.07 29.13 N/A 1073741824 8388608 float -1 35654.90 30.11 29.17 0 35691.17 30.08 29.14 N/A 2147483648 16777216 float -1 71082.79 30.21 29.27 0 70846.97 30.31 29.36 N/A 4294967296 33554432 float -1 140955.46 30.47 29.52 0 140844.77 30.49 29.54 N/A 8589934592 67108864 float -1 280461.07 30.63 29.67 0 280641.01 30.61 29.65 N/A # Out of bounds values : 0 OK # Avg bus bandwidth : 13.362 </pre>
Test Conclusion	Meets expectations
Test Notes	Pre-tuning parameters are used; the bus bandwidth at an 8G message size is taken as the test result.

4.3 AsterFabric Matrix CCL Test

4.3.1 AsterFabric Matrix All-Reduce Test

Test Case	AsterFabric Matrix All-Reduce Test
Test Objective	Test the performance of the All-Reduce communication model under the AsterFabric Matrix architecture.
Test Topology	See 2.1.4
Test Procedure	1. Build the network according to the AsterFabric Matrix architecture.

Test Result

Test Case	AsterFabric Matrix All-to-All Test
Test Objective	Test the performance of the All-to-All communication model under the AsterFabric Matrix architecture.
Test Topology	See 2.1.4
Test	1. Build the network according to the AsterFabric Matrix architecture.

Test Result

4.4.1 AsterFabric Matrix All-Reduce Tuning Test

Test Case	AsterFabric Matrix All-Reduce Tuning Test
Test Objective	Test the performance of the All-Reduce communication model under the AsterFabric Matrix architecture after parameter optimization.
Test	See 2.1.4

Topology	
Test Procedure	1. Build the network according to the AsterFabric Matrix architecture. 2. After parameter tuning, use the CCL script to test All-Reduce performance; see Expected Result 1.
Expected Result	1. Record the out-of-place busbw and verify whether performance has improved.
Test Result	1. After parameter optimization, the bus bandwidth of All-Reduce at 8G size in the Matrix network is 162.78 GB/s. <pre> # Using devices # Rank 0 Pid 1757224 on CX7Group1-host-061 device 0 [0x2b] # Rank 1 Pid 1757225 on CX7Group1-host-061 device 1 [0x3a] # Rank 2 Pid 1757226 on CX7Group1-host-061 device 2 [0x4d] # Rank 3 Pid 1757227 on CX7Group1-host-061 device 3 [0x5c] # Rank 4 Pid 1757228 on CX7Group1-host-061 device 4 [0xaa] # Rank 5 Pid 1757229 on CX7Group1-host-061 device 5 [0xba] # Rank 6 Pid 1757230 on CX7Group1-host-061 device 6 [0xca] # Rank 7 Pid 1757231 on CX7Group1-host-061 device 7 [0xda] # Rank 8 Pid 4187375 on CX7Group1-host-062 device 0 [0x2b] # Rank 9 Pid 4187376 on CX7Group1-host-062 device 1 [0x3a] # Rank 10 Pid 4187377 on CX7Group1-host-062 device 2 [0x4d] # Rank 11 Pid 4187378 on CX7Group1-host-062 device 3 [0x5c] # Rank 12 Pid 4187379 on CX7Group1-host-062 device 4 [0xaa] # Rank 13 Pid 4187380 on CX7Group1-host-062 device 5 [0xba] # Rank 14 Pid 4187381 on CX7Group1-host-062 device 6 [0xca] # Rank 15 Pid 4187382 on CX7Group1-host-062 device 7 [0xda] # Rank 16 Pid 1839036 on CX7Group1-host-063 device 0 [0x2b] # Rank 17 Pid 1839037 on CX7Group1-host-063 device 1 [0x3a] # Rank 18 Pid 1839038 on CX7Group1-host-063 device 2 [0x4d] # Rank 19 Pid 1839039 on CX7Group1-host-063 device 3 [0x5c] # Rank 20 Pid 1839040 on CX7Group1-host-063 device 4 [0xaa] # Rank 21 Pid 1839041 on CX7Group1-host-063 device 5 [0xba] # Rank 22 Pid 1839042 on CX7Group1-host-063 device 6 [0xca] # Rank 23 Pid 1839043 on CX7Group1-host-063 device 7 [0xda] # Rank 24 Pid 1716061 on CX7Group1-host-064 device 0 [0x2b] # Rank 25 Pid 1716062 on CX7Group1-host-064 device 1 [0x3a] # Rank 26 Pid 1716063 on CX7Group1-host-064 device 2 [0x4d] # Rank 27 Pid 1716064 on CX7Group1-host-064 device 3 [0x5c] # Rank 28 Pid 1716065 on CX7Group1-host-064 device 4 [0xaa] # Rank 29 Pid 1716066 on CX7Group1-host-064 device 5 [0xba] # Rank 30 Pid 1716067 on CX7Group1-host-064 device 6 [0xca] # Rank 31 Pid 1716068 on CX7Group1-host-064 device 7 [0xda] # # # size count type redop root time algbw busbw #wrong time algbw busbw #wrong # (B) (elements) # # 1024 256 float sum -1 68.89 0.01 0.03 0 65.26 0.02 0.03 0 # 2048 512 float sum -1 68.70 0.03 0.06 0 68.57 0.03 0.06 0 # 4096 1024 float sum -1 71.15 0.06 0.11 0 70.46 0.06 0.11 0 # 8192 2048 float sum -1 73.75 0.11 0.22 0 149.35 0.05 0.11 0 # 16384 4096 float sum -1 73.55 0.22 0.43 0 73.47 0.22 0.43 0 # 32768 8192 float sum -1 76.72 0.43 0.83 0 74.33 0.44 0.85 0 # 65536 16384 float sum -1 80.60 0.81 1.57 0 75.19 0.87 1.69 0 # 131072 32768 float sum -1 160.44 0.82 1.58 0 159.73 0.82 1.59 0 # 262144 65536 float sum -1 164.39 1.59 3.09 0 164.38 1.59 3.09 0 # 524288 131072 float sum -1 176.04 2.98 5.77 0 176.32 2.97 5.76 0 # 1048576 262144 float sum -1 214.46 4.89 9.47 0 205.31 5.11 9.90 0 # 2097152 524288 float sum -1 296.31 7.08 13.71 0 258.54 8.11 15.72 0 # 4194304 1048576 float sum -1 309.60 13.55 26.25 0 309.18 13.57 26.28 0 # 8388608 2097152 float sum -1 402.72 20.83 40.36 0 401.91 20.87 40.44 0 # 16777216 4194304 float sum -1 568.58 29.51 57.17 0 566.85 29.60 57.34 0 # 33554432 8388608 float sum -1 931.32 36.03 69.81 0 935.33 35.87 69.51 0 # 67108864 16777216 float sum -1 1496.73 44.84 86.87 0 1501.31 44.70 86.61 0 # 134217728 33554432 float sum -1 3394.65 39.54 76.60 0 3393.11 39.56 76.64 0 # 268435456 67108864 float sum -1 6290.76 42.67 82.68 0 6299.63 42.61 82.56 0 # 536870912 134217728 float sum -1 8219.49 65.32 126.55 0 8204.37 65.44 126.78 0 # 1073741824 268435456 float sum -1 15234.17 70.48 136.56 0 15220.77 70.54 136.68 0 # 2147483648 536870912 float sum -1 28519.77 75.30 145.89 0 28551.67 75.21 145.73 0 # 4294967296 1073741824 float sum -1 53689.93 80.00 154.99 0 53613.20 80.11 155.21 0 # 8589934592 2147483648 float sum -1 102239.69 84.02 162.78 0 102256.15 84.00 162.76 0 # 17179869184 4294967296 float sum -1 202349.67 84.90 164.50 0 202296.40 84.92 164.54 0 # Out of bounds values : 0 OK # Avg bus bandwidth : 54.766 </pre>
Test Conclusion	Meets expectations
Test Notes	Post-tuning parameters are used; the bus bandwidth at an 8G message size is taken as the test result.

4.4.2 AsterFabric Matrix All-to-All Tuning Test

Test Case	AsterFabric Matrix All-to-All Tuning Test
------------------	---

1. After parameter optimization, the bus bandwidth of All-to-All at 8G size in the Matrix network is 30.11 GB/s.

```
# Using devices
# Rank 0 Pid 1758847 on CX7Group1-host-061 device 0 [0x2b]
# Rank 1 Pid 1758848 on CX7Group1-host-061 device 1 [0x3a]
# Rank 2 Pid 1758849 on CX7Group1-host-061 device 2 [0x4d]
# Rank 3 Pid 1758850 on CX7Group1-host-061 device 3 [0x5c]
# Rank 4 Pid 1758851 on CX7Group1-host-061 device 4 [0xaa]
# Rank 5 Pid 1758852 on CX7Group1-host-061 device 5 [0xba]
# Rank 6 Pid 1758853 on CX7Group1-host-061 device 6 [0xca]
# Rank 7 Pid 1758854 on CX7Group1-host-061 device 7 [0xda]
# Rank 8 Pid 4191034 on CX7Group1-host-062 device 0 [0x2b]
# Rank 9 Pid 4191035 on CX7Group1-host-062 device 1 [0x3a]
# Rank 10 Pid 4191036 on CX7Group1-host-062 device 2 [0x4d]
# Rank 11 Pid 4191037 on CX7Group1-host-062 device 3 [0x5c]
# Rank 12 Pid 4191038 on CX7Group1-host-062 device 4 [0xaa]
# Rank 13 Pid 4191039 on CX7Group1-host-062 device 5 [0xba]
# Rank 14 Pid 4191040 on CX7Group1-host-062 device 6 [0xca]
# Rank 15 Pid 4191041 on CX7Group1-host-062 device 7 [0xda]
# Rank 16 Pid 1840439 on CX7Group1-host-063 device 0 [0x2b]
# Rank 17 Pid 1840440 on CX7Group1-host-063 device 1 [0x3a]
# Rank 18 Pid 1840441 on CX7Group1-host-063 device 2 [0x4d]
# Rank 19 Pid 1840442 on CX7Group1-host-063 device 3 [0x5c]
# Rank 20 Pid 1840443 on CX7Group1-host-063 device 4 [0xaa]
# Rank 21 Pid 1840444 on CX7Group1-host-063 device 5 [0xba]
# Rank 22 Pid 1840445 on CX7Group1-host-063 device 6 [0xca]
# Rank 23 Pid 1840446 on CX7Group1-host-063 device 7 [0xda]
# Rank 24 Pid 1717758 on CX7Group1-host-064 device 0 [0x2b]
# Rank 25 Pid 1717759 on CX7Group1-host-064 device 1 [0x3a]
# Rank 26 Pid 1717760 on CX7Group1-host-064 device 2 [0x4d]
# Rank 27 Pid 1717761 on CX7Group1-host-064 device 3 [0x5c]
# Rank 28 Pid 1717762 on CX7Group1-host-064 device 4 [0xaa]
# Rank 29 Pid 1717763 on CX7Group1-host-064 device 5 [0xba]
# Rank 30 Pid 1717764 on CX7Group1-host-064 device 6 [0xca]
# Rank 31 Pid 1717765 on CX7Group1-host-064 device 7 [0xda]
```

#	size (B)	count (elements)	type	redop	root	time (us)	out-of-place algbw (GB/s)	busbw (GB/s)	#wrong	time (us)	in-place algbw (GB/s)	busbw (GB/s)	#wrong
	1024	8	float	-1		126.46	0.01	0.01	0	131.45	0.01	0.01	N/A
	2048	16	float	-1		123.17	0.02	0.02	0	122.43	0.02	0.02	N/A
	4096	32	float	-1		121.16	0.03	0.03	0	121.12	0.03	0.03	N/A
	8192	64	float	-1		120.94	0.07	0.07	0	120.47	0.07	0.07	N/A
	16384	128	float	-1		121.92	0.13	0.13	0	120.62	0.14	0.13	N/A
	32768	256	float	-1		122.97	0.27	0.26	0	120.48	0.27	0.26	N/A
	65536	512	float	-1		127.14	0.52	0.50	0	125.12	0.52	0.51	N/A
	131072	1024	float	-1		124.99	1.05	1.02	0	124.51	1.05	1.02	N/A
	262144	2048	float	-1		133.14	1.97	1.91	0	134.41	1.95	1.89	N/A
	524288	4096	float	-1		148.73	3.53	3.42	0	148.65	3.53	3.42	N/A
	1048576	8192	float	-1		176.89	5.93	5.74	0	177.78	5.90	5.71	N/A
	2097152	16384	float	-1		210.17	9.98	9.67	0	209.25	10.02	9.71	N/A
	4194384	32768	float	-1		302.41	13.87	13.44	0	304.13	13.79	13.36	N/A
	8388608	65536	float	-1		442.28	19.97	18.37	0	441.59	19.00	18.40	N/A
	16777216	131072	float	-1		748.19	22.42	21.72	0	746.79	22.47	21.76	N/A
	33554432	262144	float	-1		1620.82	20.60	19.96	0	2726.95	12.30	11.92	N/A
	67108864	524288	float	-1		2647.19	25.35	24.56	0	2722.33	24.65	23.88	N/A
	134217728	1048576	float	-1		6104.19	21.99	21.30	0	4828.02	27.80	26.93	N/A
	268435456	2097152	float	-1		9092.57	29.52	28.60	0	9089.32	29.53	28.61	N/A
	536879912	4194384	float	-1		17921.78	39.96	29.02	0	17972.61	29.87	28.94	N/A
	1073741824	8388608	float	-1		35236.66	30.47	29.52	0	35268.45	30.44	29.49	N/A
	2147483648	16777216	float	-1		70905.28	30.29	29.34	0	70168.49	30.60	29.65	N/A
	4294967296	33554432	float	-1		139656.91	30.75	29.79	0	138875.41	30.93	29.96	N/A
	8589934592	67108864	float	-1		276490.45	31.08	30.11	0	276953.95	31.02	30.05	N/A
	17179869184	134217728	float	-1		552734.94	31.08	30.11	0	551035.47	31.18	30.20	N/A

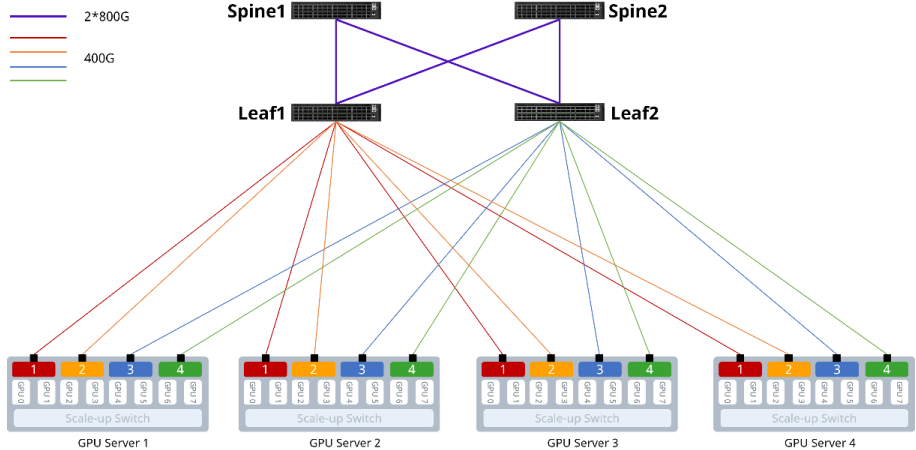
```
# Out of bounds values : 0 OK
# Avg bus bandwidth    : 13.8906
```

Meets expectations

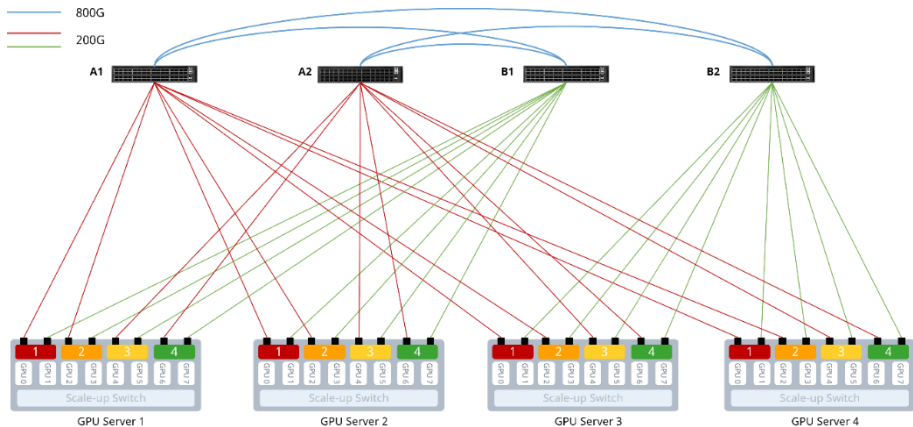
Post-tuning parameters are used; the bus bandwidth at an 8G message size is taken as the test result.

5 Model Inference Performance Test

5.1 AsterFabric ROFT DeepSeek Model Inference Test

Test Case	AsterFabric ROFT DeepSeek Model Inference Test
Test Objective	Test the inference throughput and request latency of the DeepSeek model deployed in 2P2D mode on the AsterFabric ROFT network.
Test Topology	 4 nodes × 8 GPUs, where nodes 1 and 2 are Prefill and nodes 3 and 4 are Decode.
Test Procedure	- Build the network according to the AsterFabric ROFT architecture and deploy the 2P2D service. - Load DeepSeek-V4-Flash (W8A8 INT8), use BF16 for the KV Cache, and apply the tuned configuration. - Set input to 3072 tokens and output to 1024 tokens, select concurrency levels of 16, 32 and 64, and record total throughput, average TTFT and average TPOT.
Expected Result	Complete the inference test at each level and record throughput and latency; the total throughput at the selected levels reaches the reference value.
Test Result	At concurrency 16, 32 and 64, the total throughput is 4,366.97, 6,714.27 and 10,765.62 token/s respectively, all reaching the reference throughput. The TTFT/TPOT at each concurrency level are 702.50/12.77, 1,327.98/15.69 and 2,403.08/19.14 ms respectively.
Test Conclusion	Meets expectations
Test Notes	-

5.2 AsterFabric Matrix DeepSeek Model Inference Test

Test Case	AsterFabric Matrix DeepSeek Model Inference Test
Test Objective	Test the inference throughput and request latency of the DeepSeek model deployed in 2P2D mode on the AsterFabric Matrix network.
Test Topology	 4 nodes × 8 GPUs, where nodes 1 and 2 are Prefill and nodes 3 and 4 are Decode.
Test Procedure	- Build the network according to the AsterFabric Matrix architecture and deploy the 2P2D service. - Load DeepSeek-V4-Flash (W8A8 INT8), use BF16 for the KV Cache, and apply the tuned configuration. - Set input to 3072 tokens and output to 1024 tokens, select concurrency levels of 16, 32 and 64, and record total throughput, average TTFT and average TPOT.
Expected Result	Complete the inference test at each level and record throughput and latency; the total throughput at the selected levels reaches the reference value.
Test Result	At concurrency 16, 32 and 64, the total throughput is 4,825.18, 8,369.02 and 13,101.82 token/s respectively, all reaching the reference throughput. The TTFT/TPOT at each concurrency level are 628.72/12.13, 853.88/13.66 and 1,215.97/17.34 ms respectively.
Test Conclusion	Meets expectations
Test Notes	-