EANTC Independent Test Report

Huawei WDM Compatibility Test

OptiX OSN 1800 II Pro

April 2022

Introduction

Huawei commissioned EANTC to validate functional, interoperability, and performance aspects of the Huawei OptiX OSN 1800 II Pro wave division multiplex (WDM) solution, focusing on the Storage Area Network (SAN) use case scenarios.

We conducted the test in Huawei lab in Cheng Du, China, in November 2021.

Huawei OptiX OSN 1800 II Pro is an optical-electrical WDM transmission device. Huawei OptiXtrans E6608 is a product of the same brand. Both devices provide the same implementation. It is designed for enterprise DCI use case scenarios—any situation where two redundant data centers are located in a region within a few kilometers distance (up to 100km).

In this test, EANTC verified the real enterprise case to ensure that the test results were consistent in this series of tests. The real enterprise case included two emulated data centers for the following use cases focused on Fibre Channel (FC). There is usually packet data traffic (Ethernet-/IP-based) and storage traffic forwarded over the wide-area link in a WDM DCI scenario.

Test Highlights

- → DCI interoperability with multiple Fibre Channel switches, including the combination of Brocade G620-G620, G620-6505, and Cisco MDS 9148S -9148S
- → Compatibility certification with three types of Fibre Channel Physical Interface (FC-PI), including FC-PI-3, FC-PI-5, and FC-PI-6[1]
- → Transparent multi-switch type forwarding between Brocade 6505 and Brocade G620
- → Stability of overnight soak testing
- → Protection against Inter-Switch Link (ISL) failure, ISL trunking and long haul link failure
- → Capacity measurement to link speeds of 8G, 10G, 16G, 32G with up to 100 km long haul connections[2]

Executive Summary

We verified the interface-speed forwarding of Huawei OptiX OSN 1800 II Pro on the FC port side and the long-distance forwarding between the two simulated data centers. We tested the forwarding performance when a quorum server synchronized data traffic between both emulated data centers. We verified the compatibility of its board types LDCA and LTX to the FC-PI-6, FC-PI-3, and FC-PI-5 standards.

The test bed consisted of emulated data centers integrated with 3rd party Fibre Channel switches of different vendors represented by Brocade and Cisco, using hybrid switch pairs of Brocade G620 / Brocade G620, Brocade G620 / Brocade 6505, and Cisco MDS 9148S/ Cisco MDS 9148S.

We verified the robustness of the DUT by performing administrative activities on the DUT and connected equipment, as well as ISL trunking and protection against long-haul link failure. We also put DUT under continuous load for 24 hours in a soak test environment to make sure it would support uninterrupted service. The system remained stable without any restart or service interruption, zero packet loss, and low latency, as expected. Finally, we measured the latency introduced by the OptiX OSN 1800 II Pro. It matched the expectations based on switching delay and physical distance.

[1] FC-PIs specifications are defined by the T11 Committee of the International Committee on Information Technology Standards (INCITS). INCITS is accredited by and operates under rules approved by the American National Standards Institute (ANSI). FC-PI-6 (ANSI INCITS 512-2015) defines the standard to support the link speeds of 32G, 16G, and 8G; FC-PI-5 (ANSI INCITS 479-2011) defines the standard to support the link speeds of 16G, 8G, and 4G; FC-PI-3 (ANSI/INCITS 460-2011) defines the standard to support the link speeds of 10G, 4G, 2G, and 1G.

[2] Cisco MDS 9148S was tested with 40 km long-haul at FC8G and no 24-hour stability test

Device Under Test

Huawei explained that the OptiX OSN 1800 II Pro is designed for DCI and ready for simplified deployment, ultrabroadband and high integration data traffic.

Figure 1: Huawei OptiXtrans 1800 II Pro

Figure 2: Huawei LDCA Board

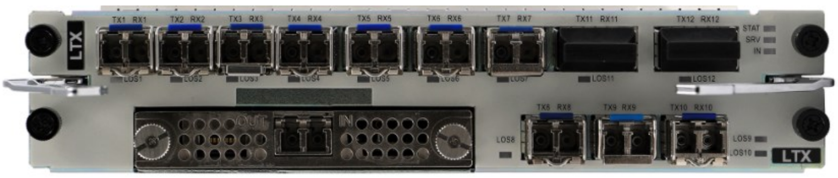

Figure 3: Huawei LTX Board

Compatibility Test Combinations

Set-up	FC Switch1	FC Switch 2	Optical Trans-ceiver (at E-port)
1	Brocade G620-1	Brocade G620-2	FC-PI-6
			FC-PI-3
2	Brocade G620-1	Brocade 6505[3]	FC-PI-5
3	Cisco MDS 9148S-1	Cisco MDS 9148S-2[4]	FC-PI-5

Table 1: FC Switch Combinations

FC Switch	Huawei OptiX OSN 1800 II Pro / Huawei OptiXtrans E6608
FC-PI-6	FC-PI-6
FC-PI-5	FC-PI-5 and FC-PI-6
FC-PI-3	FC-PI-3

Table 2: Optical Transceivers between FC Switch and DUT

Long haul connection tests were successfully carried out across two distances:

- 100 km long haul with Brocade G620 / G620 and Brocade G620 / 6505 pairs, respectively
- 40 km[5] calculated medium-haul with Cisco MDS 9148S pairs, respectively

[3] Brocade 6505 does not support FC-PI-6 and FC-PI-3

[4] Cisco MDS 9148S does not support FC-PI-3 and FC-PI-6

[5] Due to Cisco MDS 9148S buffer limitation, we can only test up to 40 km.

Testbed Description

The Huawei OptiX OSN 1800 II Pro and OptiXtrans E6608 were the device under test (DUTs) respectively. They have the same hardware, and only the product name is different for a different market. Therefore, we designed the test bed as shown in the figure below.

Huawei installed SUSE Enterprise Linux 12.4 in a bare metal mode on both hosts. We used vdbench already installed at Huawei labs by Huawei engineers to generate FC traffic. The host has 2x Intel® Xeon® E5-2658 v4 @2.30GHz CPUs, 8x 16G DDR4 memory, 1 x Huawei IN300 (2x FC32G) port FC Host Adapter (HBA), 1 x Emulex LPe16002B-M6 PCIe 2-port 16Gb Fibre Channel Adapter and 1x SAS 800G SSD. We used all the CPU, FC adapter ports, and 16G memory for our test.

The storage hardware included two Huawei OceanStor 5000 V3 (referred to as Huawei OceanStor) devices, each equipped with 24 Non-Volatile Memory Express (NVMe) Solid-State Drive (SSD) disks that provided up to 5.8 GB/s (Gigabytes per second) Input/Output traffic. Using the open source test tool Vdbench released by Oracle, we generated bidirectional baseline traffic at the full configured speed.

The key point from the topology was the quorum server in place. It means that the storage system also had redundancy control. The quorum server recognized one of the Huawei OceanStor 5000 V3 as preferred storage and another one as non-preferred storage. The quorum server's strategy was to keep the preferred storage in a working state when a link or storage failure is detected.

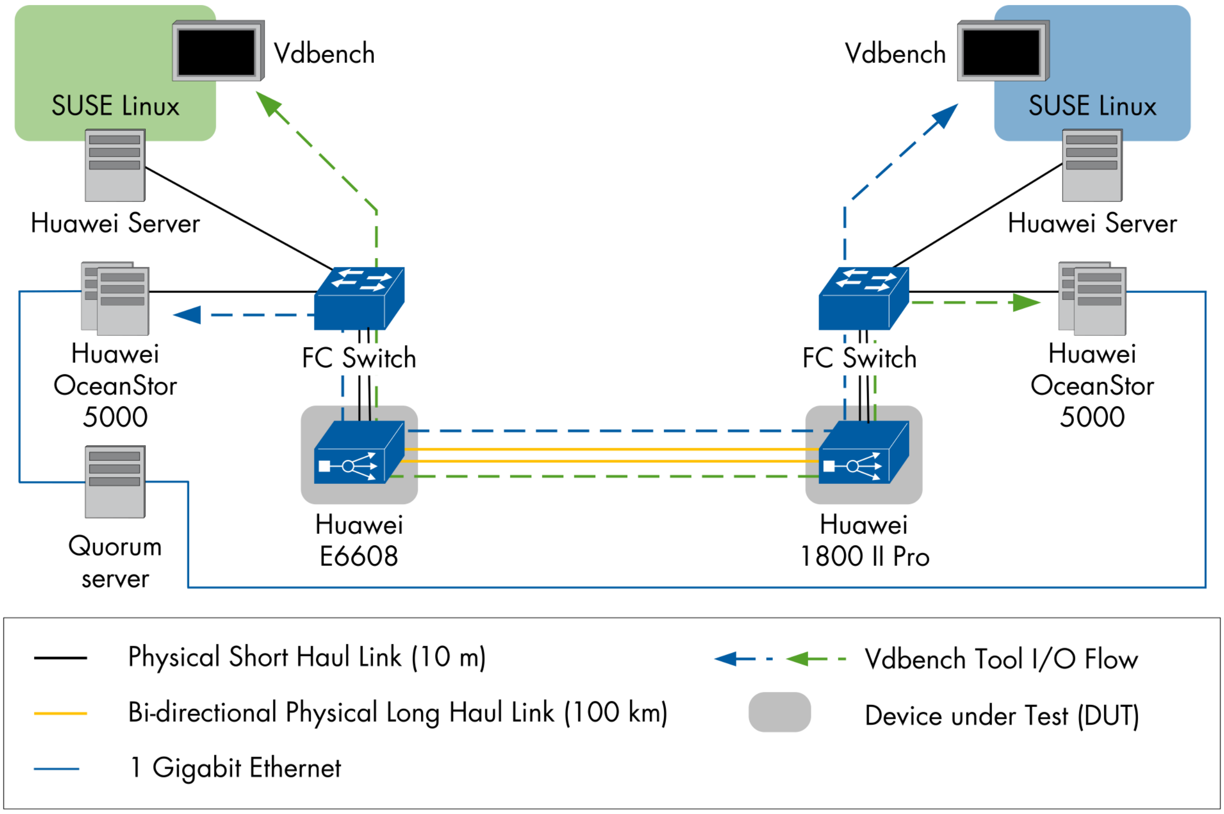

Figure 2: Logical Data Center Setup

Hardware and Software

Product Type	Product Name	Software Version
Devices Under Test WDM Equipment	Huawei OptiX OSN 1800 II Pro / Huawei OptiXtrans E6608	V100R021C00
	LDCA (board facing FC switch)	
	LTX (board facing FC switch)	
	OBU	
	OLP (board facing long-haul)	
	UXCL SCC STG	
Physical Server	Huawei RH2288H V3	iBMC: 2.94 BIOS: 3.87
Operating System	SUSE Linux Enterprise Server 12 SP4	Release: 12.4 Kernal: Linux 4.12.14-94.41-default
FC traffic simulation software	Vdbench	v50406
FC traffic generator	Viavi MTS5800-100G	BERT 28.0.1
SAN Storage	Huawei OceanStor 5000 V3[6]	V300R002C10
Quorum server	Quorum server	V300R002C10

Table 3: Hardware and Software Components

FC Switch	Software Info
Brocade 6505	FOS v8.2.1a
Brocade G620	FOS v9.0.1c
Cisco MDS 9148S	NX-OS version 6.2

Table 4: FC Switch Hardware and Software Details

[6] 24 Non-Volatile Memory Express (NVMe) solid-state disk (SSD) disks that provided up to 5.8 GB/s (Gigabytes per second) Input/Output

Test Results

Capacity Testing

We measured the maximum FC-interface forwarding of the DUT using FC read block traffic (reading block size 32KiB) generated respectively between emulated data centers.

The following tables show the throughput. We performed a capacity test for each of the board under test, with all three setups, respectively. Each traffic stream carried bi-directional traffic.

Setup	Speed Type	Expected Throughput (MB/s)[7]	Measured Throughput, per direction (MB/s)		Verdict
			LDCA	LTX	
1. Brocade G620 pair	FC800	800	781	781	Pass
	FC1200	1200	1172	1168-1171	Pass
	FC1600	1600	1558	1560-1561	Pass
	FC3200	3200	3113-3117	3113-3115	Pass
2. Brocade G620/6505 pair	FC1600	1600	6505: 1529 G620: 1558	6505: 1531 G620: 1563	Pass
3. Cisco MDS 9148S/9148S	FC800	800	778	777	Pass

Table 5: LDCA and LTX Interface Throughput

[7] The expected throughput is based on the layer 2 payload; we used a ratio of 97% of the link speed. The ratio is based on the formula: 2,048 bytes payload size / 2,112 bytes maximum frame size * 100%, which excludes overhead from the throughput consisting of Start Of Frame (SOF), Cyclic Redundancy Check (CRC), and End of Frame (EOF). For example, 3.104 GB/s = 97% * 3.2 GB/s link speed.

Latency Test

We measured with the Viavi MTS5800-100G tester the WDM system's latency value. In this setup, we connected both DUTs back to back with the measurement tool (see figure), which provides latency value in microseconds' precision. We removed all data center devices from the test bed and remained only both WDM devices running the FC services between the traffic generators.

We generated FC traffic from all ports at different FC speeds consisting of 8G, 10G, 16G, and 32G. The test tool supported the latency measurement on the same port. Therefore, we designed the Rx and Tx like below. Using an optical splitter, we split the Rx and Tx at the traffic generator port into two separate fibers and connect them to the two WDM devices as shown in the figure.

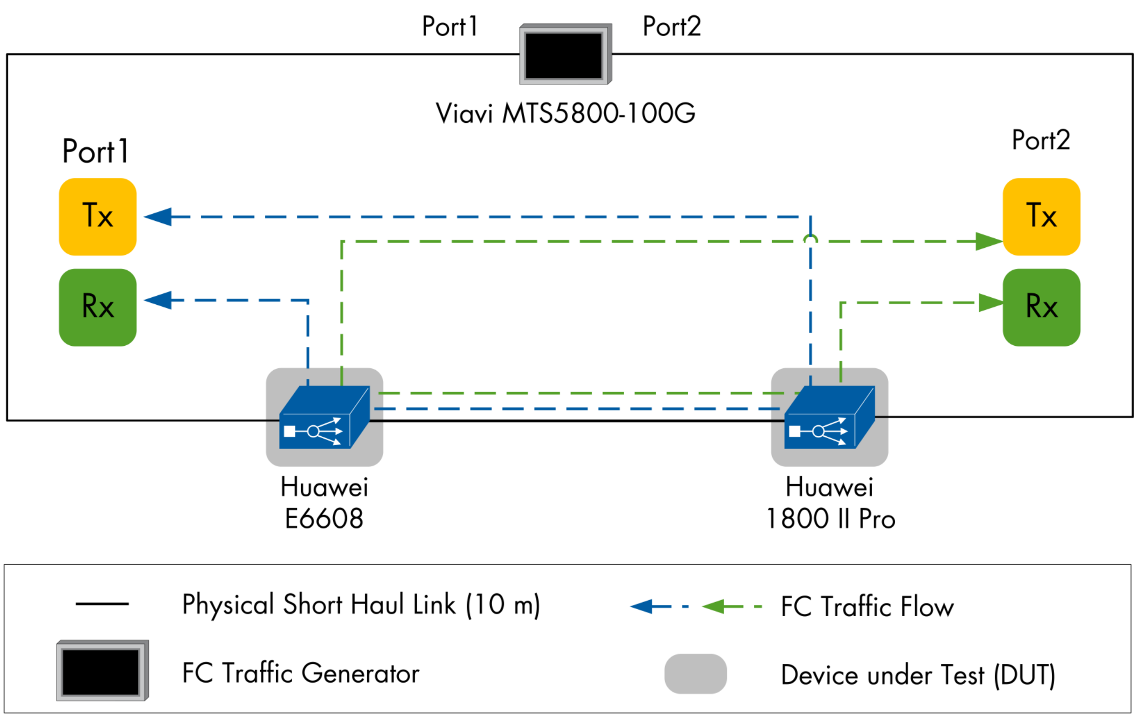

Figure 3: Latency Test Topology

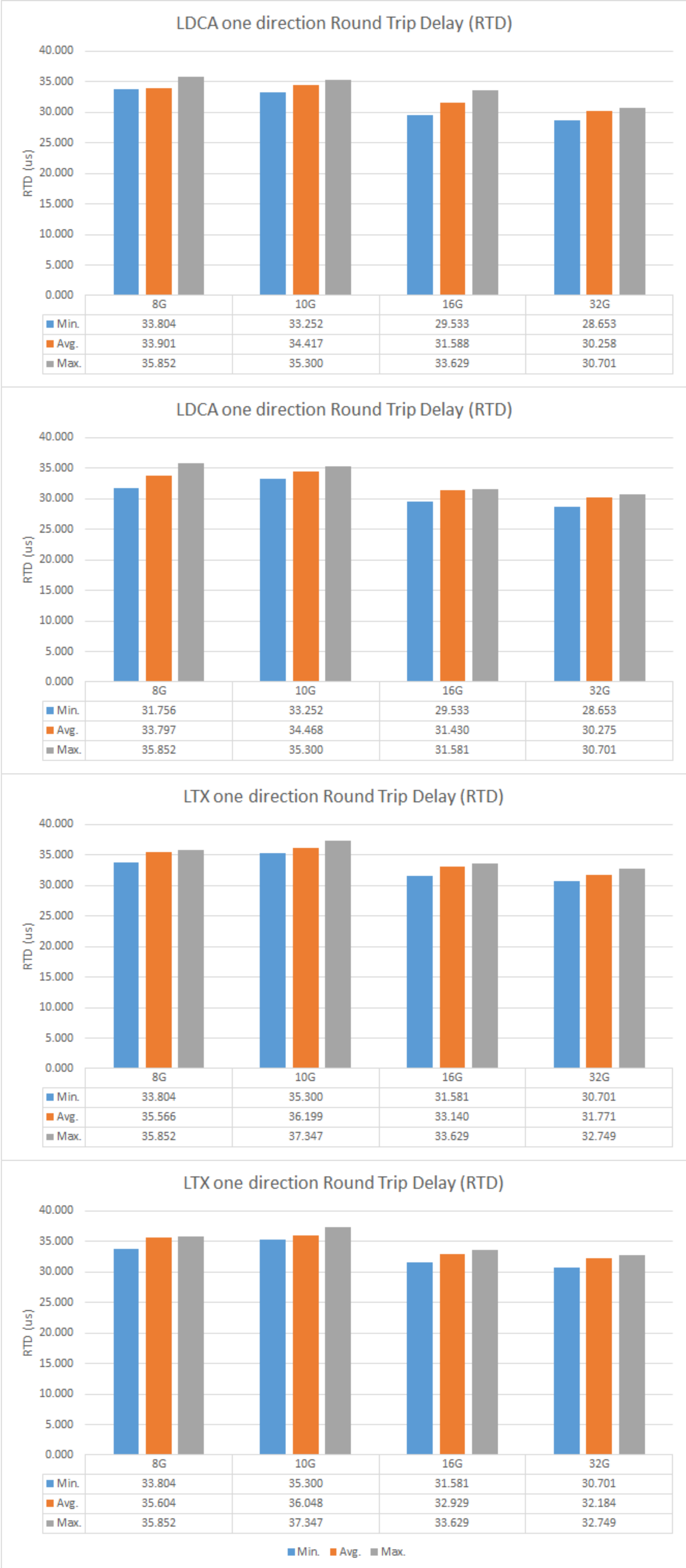

Figure 4: Latency Results of LDCA and LTX

Availability Test

We verified that the DUT maintained stability when performing administrative activities on the DUT's hardware and the FC switch connected in the test environment.

We emulated a baseline scenario where the host's I/O traffic flowed between SANs from two different emulated data centers under normal conditions. The host on each side of the DCI initiated I/O operations, and the target was the storage on the other side of the DCI. The following figure depicts the main traffic streams (blue and green, both bidirectional) between the emulated data centers.

To verify the stability of the DUT, we emulated a common set of failures of the DCI segment while the traffic was running under normal conditions. These were a total of 7 types of emulated failures as shown in the figure.

In ISL trunking and long-haul link protection, we expected the DUT to provide traffic switching between primary and redundant links. Especially with the participation of Quorum servers, we expected that the active-active storage cluster to protect the traffic switching between storage devices.

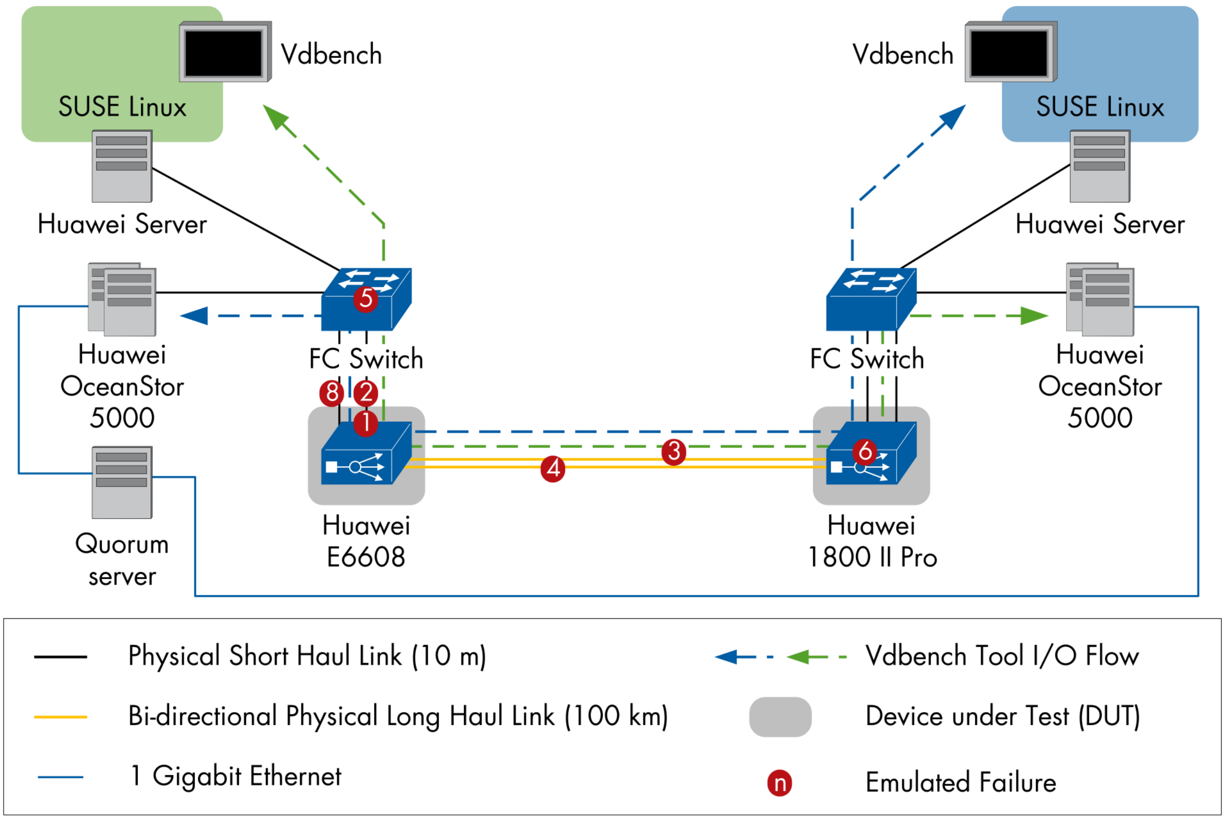

Figure 5[8]: Emulated Failure in the Test Bed

8 The numbers in figure 5 represent the test cases in the test plan. Missing of number 7 is because 7 represents 24 hour long-haul test, and it doesn't include any failure scenario.

Active-Active Storage Switch Over

Under failure types 1, 2, 3, and 6, the test results showed the same status of the traffic that has been switched over. The primary storage obtained the I/O access of the host in the same data center and continued I/O operation. The secondary storage stopped receiving any I/O access from the remote datacenter, neither from the local data center. The former case was because these four types of failure could interrupt the whole DCI connection from different hardware locations (see figure below), thereby causing it to stop transmitting data between both data centers. The latter phenomenon was due to the design mechanism of quorum server as described below.

Huawei OceanStor 5000 V3 implements a technology called "HyperMetro" for synchronizing the states between preferred storage and non-preferred storage. Once the link between two storage systems went down, the HyperMetro pair changed to the "To be synchronized" state. The Logic Unit Number (LUN) in the preferred storage continued providing service while the LUN in non-preferred storage stopped. During a link recovery, once the quiescing time (300 s) passed, we observed that the traffic switched back to baseline status between DC1 and DC2. We did not observe any impact as expected.

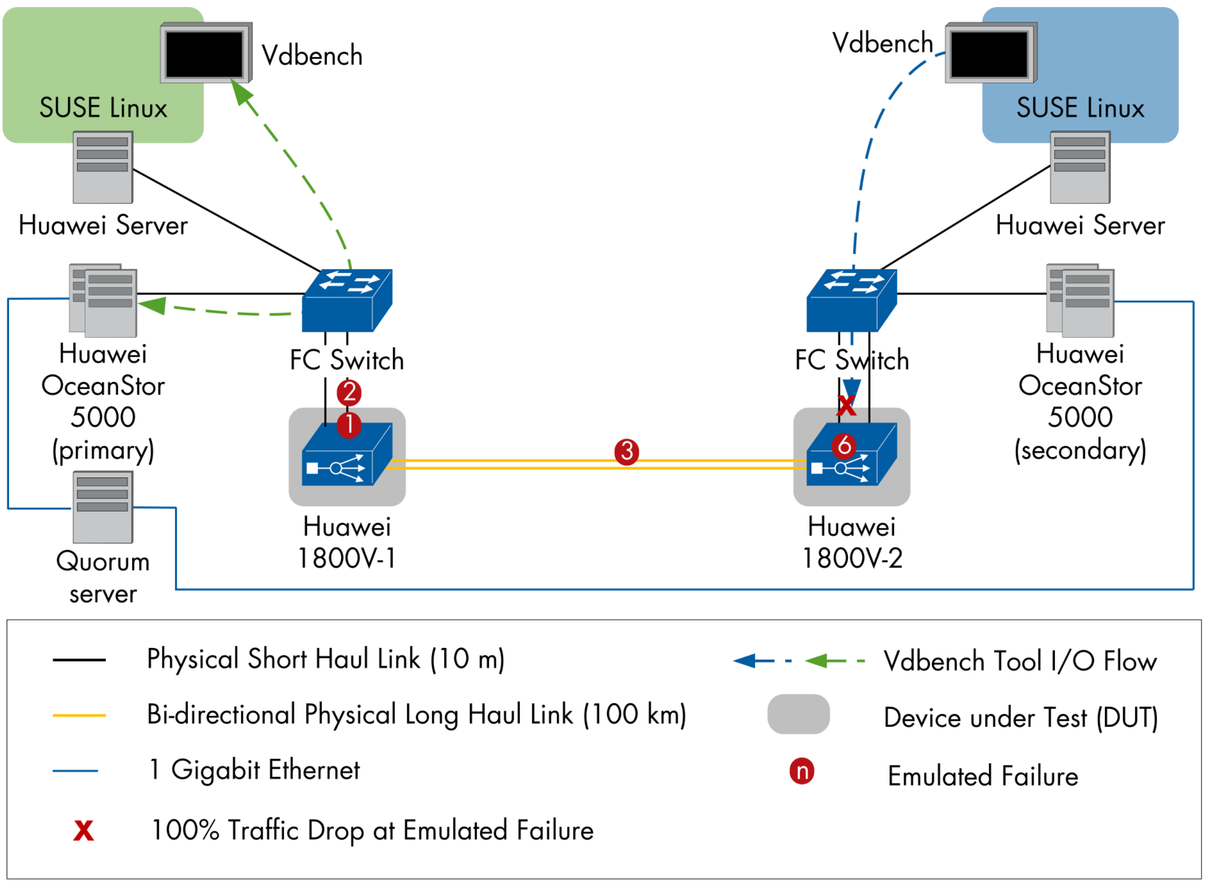

Figure 6: Traffic Status after the Switch Over for Failure Types 1, 2, 3 and 6

No. in Fig. 6	Test Case	Action	Service Interruption		Verdict
			Expected[9]	Observed (for primary storage)[10]	
1	E-Port Disable/ Enable Test	Disable	10s	A maximum of 7s service interruption time, including 5s drop to 0MB/s	Pass
		Enable	No impact	No impact	
2	E-Port Cable Disconnect/ Reconnect Test	Disconnect	10s	A maximum of 7s service interruption time, including 5s drop to 0MB/s	Pass
		Reconnect	No impact	No impact	
3	Long Haul Network Failure without Redundancy	Disconnect	10s	A maximum of 7s service interruption time, including 5s drop to 0MB/s	Pass
		Reconnect	No impact	No impact	
6	DUT Reset Test	Shut down	10s	A maximum of 7s service interruption time, including 5s drop to 0MB/s	Pass
		Turn on	No impact	No impact	

Table 6: Out of Service Overview (Active-Active Storage Cluster Protection)

[9] The expected value included 5s fault detection by the quorum server and 15s from it to triggering an active-active switchover. Once the quorum server detected the fault, it triggers an active-active switchover, which lasts for 15 seconds (generally within 10 seconds in lab tests). After the switchover has been complete, the preferred storage took over services.

[10] We observed 5s complete traffic drop to 0 MB/s during the switch over. The Huawei team configured the quorum server with a 5s timer to detect the link heartbeat failure between the storage arrays. The link between storage arrays sent a heartbeat packet every second. After five consecutive heartbeat packets expired, the link was identified as disconnected. The 6s complete traffic drop to 0 MB/s included the time from discovery to switchover at the quorum server.

FC Switch Reboot

Once the FC switch from one data center rebooted, the storage on the remote site where the FC switch was not touched obtained the host's I/O access and continues to perform I/O operations. In the data center where the FC switch rebooted, the storage stopped receiving any data from the FC switch until the FC switch was available again. Depending on the time taken by the FC switch to become available, we observed:

- The Brocade G620 pair rebooted less than 1 min. The FC service did not fully interrupt from the reboot site. The reboot site traffic was redirected to the non-reboot site and kept working, whereas the non-reboot site switched to access the local storage as well. There was around 40-50 s service interruption time (including a maximum of 39s drop to 0MB/s).

After the FC switch completed the reboot and the quiescing time (300s) passed, we observed that the traffic switched back to the initial status between DC1 and DC2. We did not observe any impact as expected.

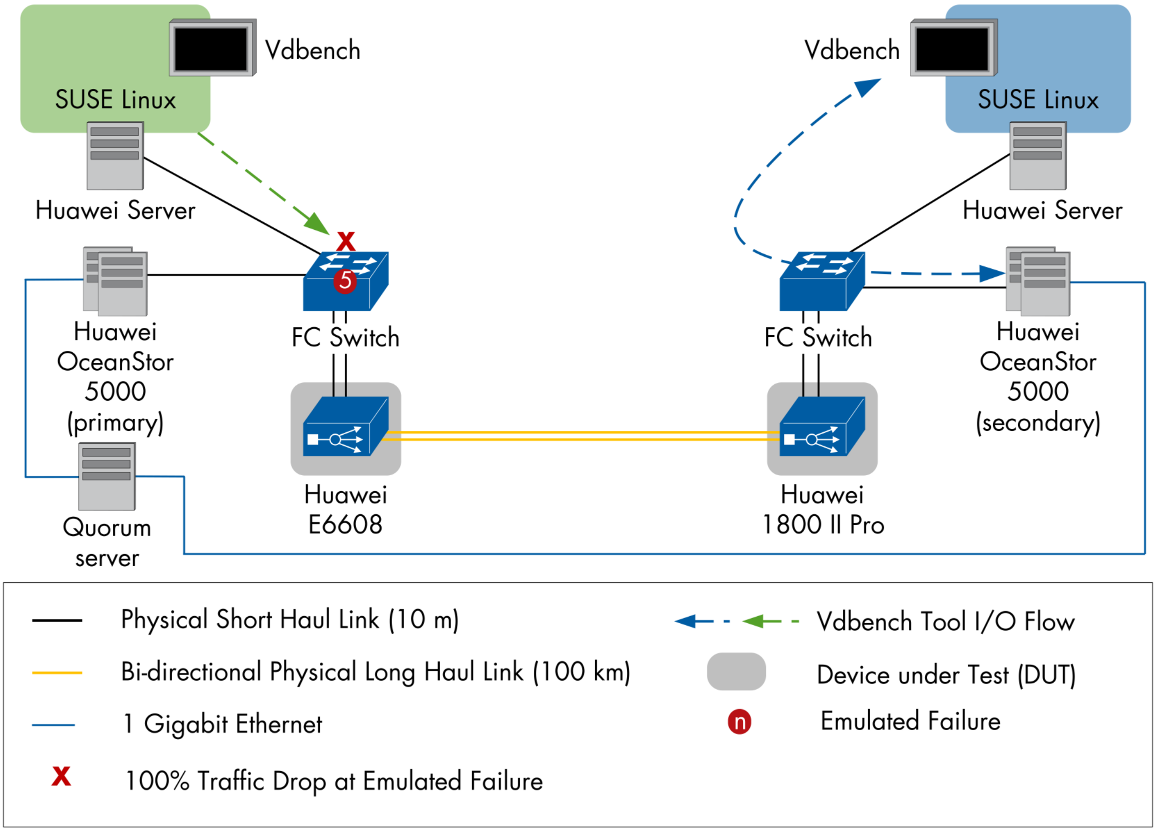

Figure 7: Traffic status during FC switch reboot

No. in Fig. 7	Test Case	Action	Service Interruption		Verdict
			Expected[11]	Observed (for primary storage)[12]	
5	Switch Reboot Test	Shut down	15s	A maximum of 10s service interruption time on non-reboot site, including a maximum of 8s drop to 0MB/s	Pass
		Turn on	No impact	No impact	

Table 7: Out of Service Overview (Active-Active Storage Cluster Protection)

[11] The expected value included 5s fault detection by the quorum server and 15s from it to triggering an active-active switchover.

[12] The 8s drop to 0MB/s included 6s of switch over by the quorum server and the reboot time of the switch (during the reboot process, the link at the FC switch was interrupted from time to time)

Long Haul Link Protection

We measured the switch-over time for the system under test to switch traffic to the backup link when the primary long-haul link fails. While traffic was running, we disconnected the primary link from the long-haul connection between both WDM devices. We observed that the traffic was switched to the backup link. The following Table shows the switch-over time measured. No session drops appeared during this time.

When we reconnected the primary link previously disconnected, we did not observe any impact on the traffic. Then we disconnect the protect link again. The WDM switched back to the primary link again with the switch over time. The last step is to reconnect the protected link. We observed no impact on I/O flow traffic.

ISL trunking is a resiliency technology that Brocade uses to improve performance and redundancy. When two E-ports have the exact same configuration and are in the same zone, these two E-ports can be seen as a bundle. It has the load-share function that the two ports can achieve double throughput of the configured speed. They are also back-up for each other when one link has a problem or is down for unknown reasons. In our test, the ISL trunking was up and running at the beginning of the test. We unplugged fiber from one of the trunking ports, and we expected the traffic to reduce to 50% but no interruption. We then plugged the fiber back to its original port, and we expect the traffic to increase by 100%.

No. in Fig. 5	Test Case	Setup	Action	Service Interruption		Verdict
				Expected[13]	Observed	
4	Long Haul Network Failure with Redundancy	Brocade G620 pair	Disconnect	4s[14]	A maximum of 4s service interruption time, including 2s drop to 0MB/s	Pass
			Reconnect	No impact	No impact	
		Brocade G620 6505 pair	Disconnect	4s[14]	A maximum of 4s service interruption time, including 1s drop to 0MB/s	Pass
			Reconnect	No impact	No impact	
8	ISL Trunking	Brocade G620 pair	Disconnect	Drop 50%	Drop 50%	Pass
			Reconnect	Increase 100%	Increase 100%	

Table 8: Out of Service Overview

[13] Includes the impact of end-to-end flow control (a total of four hops from the host to the storage through two FC switches). With credit recovery enabled on the FC switch, we calculated each hop for 1 second interruption, based on the hold off time of 500ms (milliseconds that a frame could be buffered on a port without being overwritten) configured on the Brocade G620 switch; added to that the impact of retransmission caused by the frame loss during the link failure, and vdbench accuracy of 1 sample per second.

[14] The expected value includes 4s of impact by end-to-end flow control (see explanation as provided in * note), plus 2s of impact by edge hold off time (EHT) on F port. Note, the value is not provided by Brocade 300 manual, we calculated a double theoretical value of 250ms for FC1G and FC2G (250ms is the default EHT value of FC16G), so added 1s per F port (based on 1 sample/sec based on vdbench).

Board Under Test	Switch Over Time (s)
LDCA	A maximum of 3s service interruption (including a maximum of 1s drop to 0MB/s)
LTX	A maximum of 3s service interruption (including a maximum of 1s drop to 0MB/s)

Table 9: Switch Over Time Brocade G620 - G620

Board Under Test	Switch Over Time (s)
LDCA	A maximum of 3s service interruption (including a maximum of 1s drop to 0MB/s)
LTX	A maximum of 3s service interruption (including a maximum of 1s drop to 0MB/s)

Table 10: Switch Over Time Brocade G620 - 6505

Summary of Test Runs for All Failure Scenarios

With all boards under test, the DUT demonstrated its ability to maintain stability while performing the above failure scenarios in each of the test setups.

We used the maximum load of the baseline traffic for each of the selected interfaces as measured in capacity tests and formed 36 combinations as listed in the tables below.

Board Under Test/ Test Scenario	E-port Interface used in Test Case						
	1	2	3	4	5	6	8
LDCA	8G, 16G	10G, 32G	8G, 32G	10G, 16G	8G, 16G	32G	16G, 32G
LTX	8G, 16G	10G, 32G	8G, 32G	10G, 16G	8G, 16G	32G	16G, 32G

Table 11: Brocade G620 Pair

Board Under Test/ Test Scenario	E-port Interface used in Test Case		
	1	4	6
LDCA	16G	16G	16G
LTX	16G	16G	16G

Table 12: Brocade G620 - 6505 Pair

Board Under Test/ Test Scenario	E-port Interface used in Test Case	
	2	6
LDCA	8G	8G
LTX	8G	8G

Table 13: Cisco MDS 9148S Pair

Soak 24 Hours Test

We verified the WDM system's reliability in terms of performance consistency under long-period stress load conditions. The Huawei team configured all three boards in a snake configuration, keeping traffic flowing between the two data centers.

The test tool Vdbench triggered the baseline traffic bidirectional for 24 hours. During that time, we monitored the system log of both hardware and software. We confirm that the system could transfer the data at a consistent rate and constant latency of 24 hours. As expected, the system under test remained stable; we did not observe any software crashes or hardware failures during the test duration.

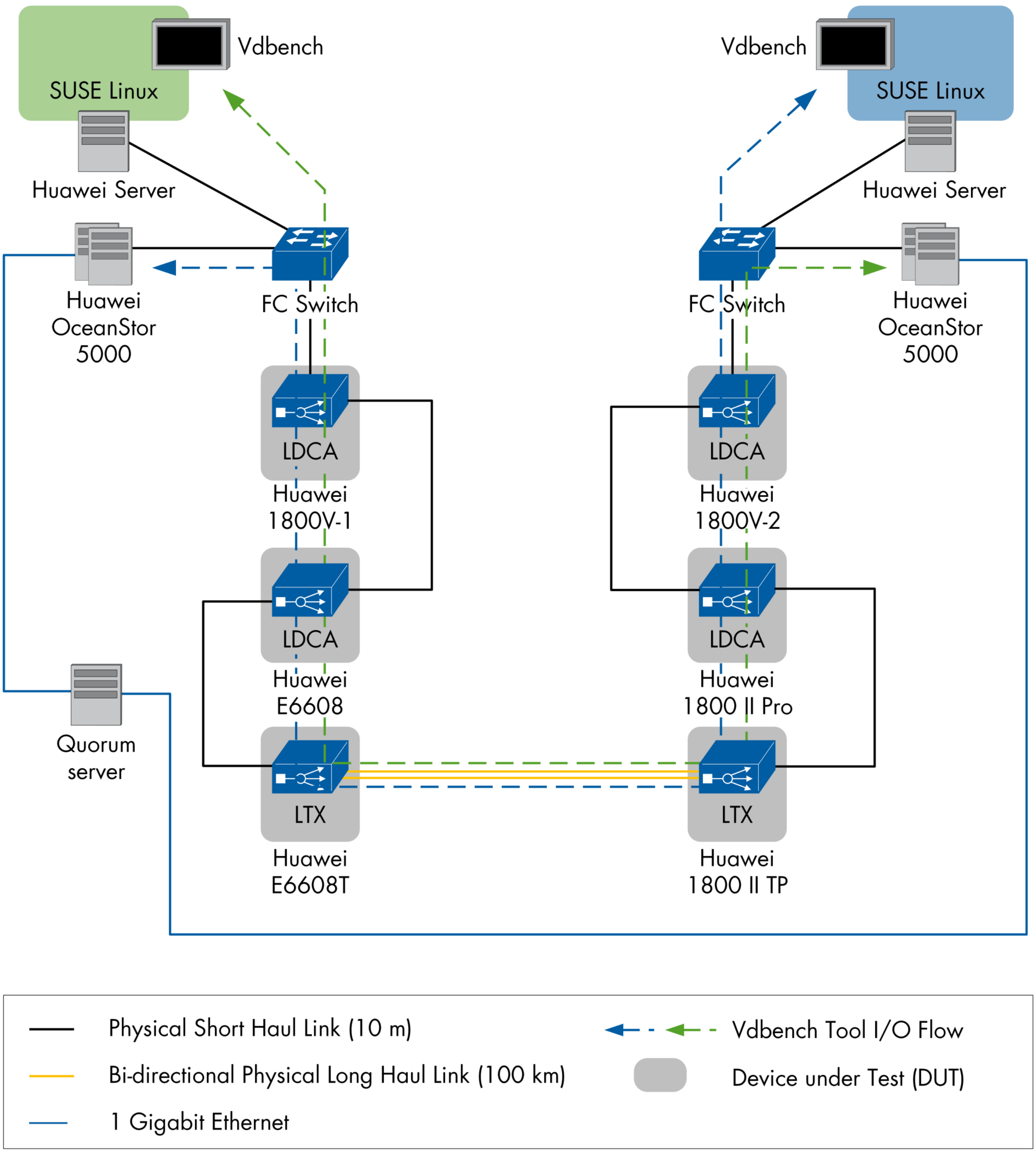

Figure 8: 24-hour Throughput Topology of LDCA

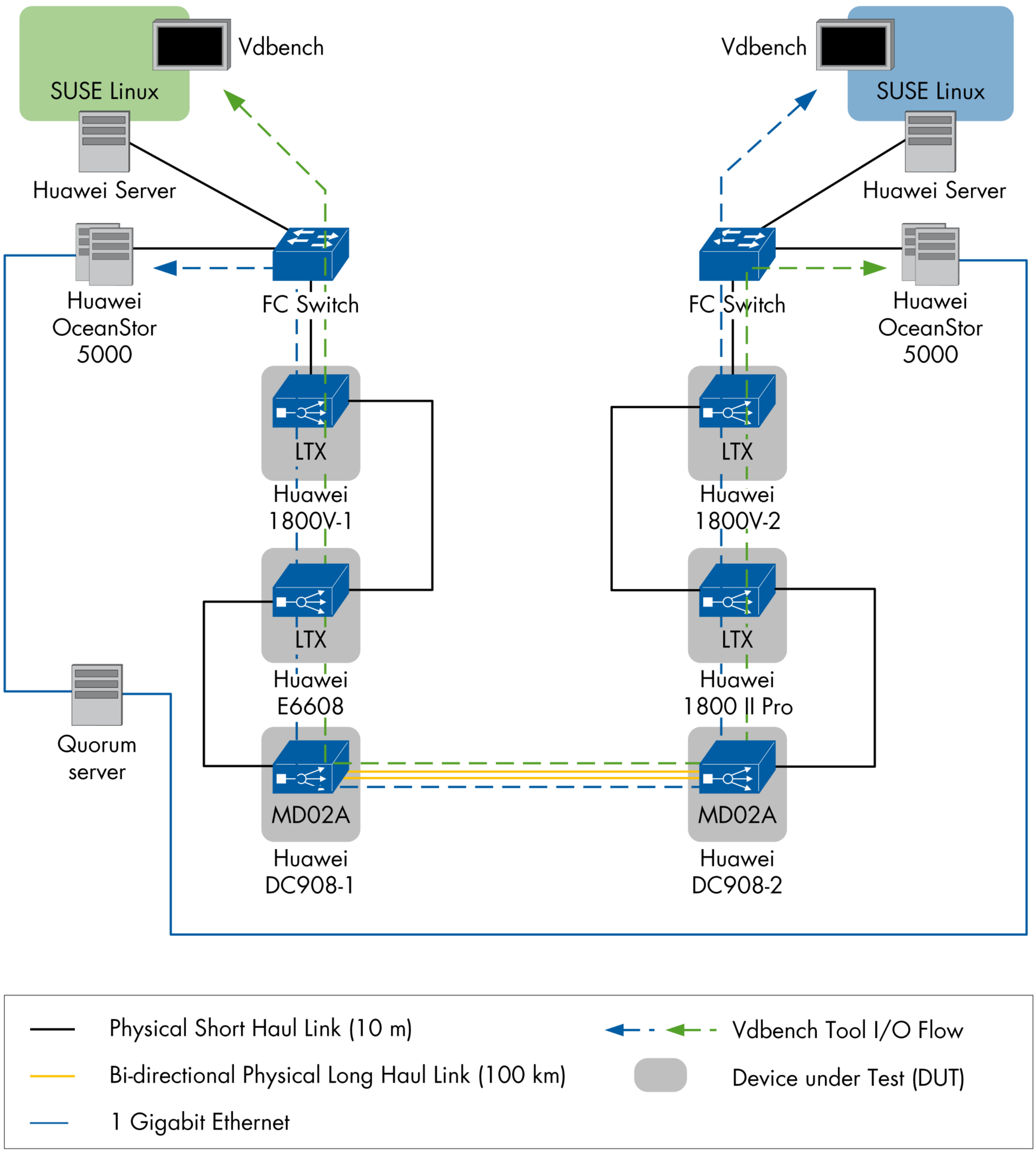

Figure 9: 24-hour Throughput Topology of LTX

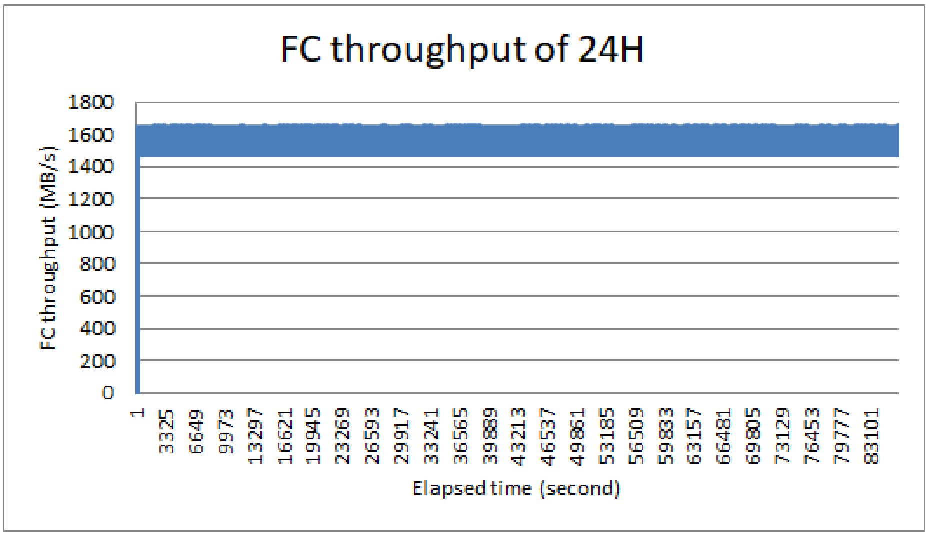

Figure 10: 24-hour Throughput of LDCA at FC 16G

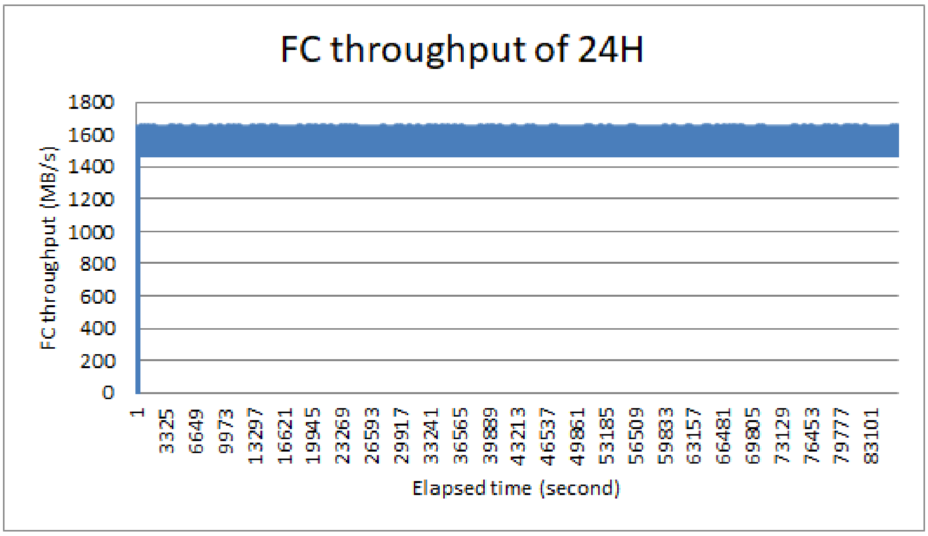

Figure 11: 24-hour Throughput of LTX at FC 16G

Conclusion

We verified interoperability of Huawei OptiX OSN 1800 II Pro and third-party Fibre Channel switches from Brocade (G620/G620, G620/6505) and Cisco (MDS 9148S/9148S). Multiple optical transceiver functions were certified in the E_port between these FC switches and the Huawei OptiX OSN 1800 II Pro, including FC-PI-6, FC-PI-5, and FC-PI-3.

EANTC validated forwarding speeds in FC 8G, 10G, 16G, and 32G scenarios. When forwarding traffic at any of these standardized speeds, the Huawei OptiX OSN 1800 II Pro did not exhibit any speed impact. The extended 24 hours soak testing confirmed stable operations of OptiX OSN 1800 II Pro without any traffic impact.

We conducted a range of availability tests to disable/enable E_port, disconnect/reconnect E_port fiber, disconnect/reconnect long haul, reboot FC switch, remove/re-install LDCA and LTX board, and remove/re-install ISL trunking fiber. All of the tests passed our verification.

Based on our test results, EANTC confirms that the Huawei OptiX OSN 1800 II Pro fulfils Huawei's claims to work in enterprise data center interconnection scenarios as an integrated, high speed, highly available solution.

EANTC AG
Salzufer 14, 10587 Berlin, Germany
info@eantc.de, https://www.eantc.de/
[v1.0 20220607]