EANTC Independent Test Report

Huawei WDM Compatibility Test

OptiXtrans DC908

April 2022

Introduction

Huawei commissioned EANTC to validate functional, interoperability, and performance aspects of the Huawei OptiXtrans DC908 wave division multiplex (WDM) solution, focusing on the Storage Area Network (SAN) use case scenarios. We conducted the test in two phases: the first test session took place in our lab in Berlin, Germany, in May 2020, targeting the Huawei V100R019C10 software version. In November 2021, the EANTC team dedicated again to the regression test of the new software version V100R021C10, and the latter test took place in the Huawei lab in Cheng Du, China.

Huawei OptiXtrans DC908 is an optical-electrical WDM transmission device. 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 100 km).

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.

Device Under Test

Huawei explained that the OptiXtrans DC908 is designed for DCI and ready for simplified deployment, ultrabroadband and high integration data traffic.

Figure 1: Huawei DC908

Test Highlights 2020

- → DCI interoperability with four Fibre Channel switches, including Brocade 6505, Brocade G620, Cisco MDS 9148S, and Cisco MDS 9132T
- → Compatibility certification with three types of Fibre Channel Physical Interface (FC-PI), including FC-PI-3, FC-PI-5, and FC-PI-6[1]
- → Capacity measurement to link speeds of 8G, 10G, 16G, 32G with 100 km long haul connections
- → Transparent multi-switch type forwarding between Brocade 6505 and Brocade G620, as well as between Cisco MDS 9148S and Cisco MDS 9132T, respectively
- → Protection against Inter-Switch Link (ISL) failure, and long haul link failure
- → Stability of overnight soak testing

Test Highlights 2021

- → DCI interoperability with multiple Fibre Channel switches, including the combination of Brocade G620-G620, G620-6505
- → 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 1G, 2G, 4G, 8G, 10G, 16G, 32G with up to 100 km long haul connections

[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.

Executive Summary

We verified the interface-speed forwarding of Huawei OptiXtrans DC908 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 MD02A 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 9132T / Cisco MDS 9132T, Cisco MDS 9132T/ Cisco MDS 9148S, respectively in 2020. We verified hybrid switch pairs of Brocade G620/ Brocade G620, Brocade G620 / Brocade 6505, and Brocade 300 / Brocade 300 in 2021.

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 DC908. It matched the expectations based on switching delay and physical distance.

This report contains two parts, the results of the DUT with the software version (V100R019C10), and the results of the new software version (V100R021C10-SPC100) to ensure comprehensiveness.

FC Switch	Huawei OptiXtrans DC908
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 and Cisco MDS 9132T pairs, respectively
- 20 km[4] calculated medium-haul with Brocade 6505 and Cisco MDS 9148S pairs, respectively

Setup	FC Switch1	FC Switch 2	Optical Transceiver (at E-port)
1	Brocade G620-1	Brocade G620-2	FC-PI-6
			FC-PI-3
2	Brocade G620-1	Brocade 6505[2]	FC-PI-5
3	Cisco MDS 9132T-1	Cisco MDS 9132T-2[3]	FC-PI-6
4	Cisco MDS 9132T-1	Cisco MDS 9148S-2	FC-PI-5

Table 1: FC Switch Combinations

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

[3] Cisco MDS 9132T does not support FC-PI-3

[4] Since none of the switches actually supported 100 km distances due to protocol settings, we calculated theoretical kilometer distance expected at the Brocade 6505 and the Cisco MDS 9148S based on the buffer supported on the switches (theoretical distance (kilometer) = (bytes of buffer size per port group - number of ports per port *1 reervd byte - 1 port) / recommended bytes of buffer size per kilometer).

2020

The following sections describe the test results of the V100R019C10.

Testbed Description

We verified the pair of the DUTs in all test cases to make sure they had the same performance/behavior. The Inter-switch link (ISL) functions between Brocade G620/G620, and Brocade G620/6505 was also in place. Huawei installed SUSE Enterprise Linux 12.4 in a bare metal mode on both hosts. We used vdbench which was already installed at Huawei labs by Huawei engineers for generating FC traffic.

The host has 2x Intel® Xeon® E5-2658 v4 @2.30GHz CPUs, 8x 16G DDR4 memory, a 2x FC32G port FC adapter, and 1x NVMe SSD. We used all the CPU, FC adapter ports, and 16G memory for our test.

Hardware and Software

The storage hardware included two Huawei OceanStor 5500 V5 (referred to as Huawei OceanStor) devices, each equipped with 20 serial attached SCSI (SAS) hard disk drive (HDD) disks that provided up to 1.6 GB/s (Gigabytes per second) Input/Output traffic. Using the open source test tool Vdbench released by Oracle, we generated bidirectional baseline traffic on the full traffic throughput for 8G, 10G and 16G FC-link. We chose the same baseline traffic at 1.6 GB/s for 32G FC-link to verify the functionality of the WDM system integrated with the data center.

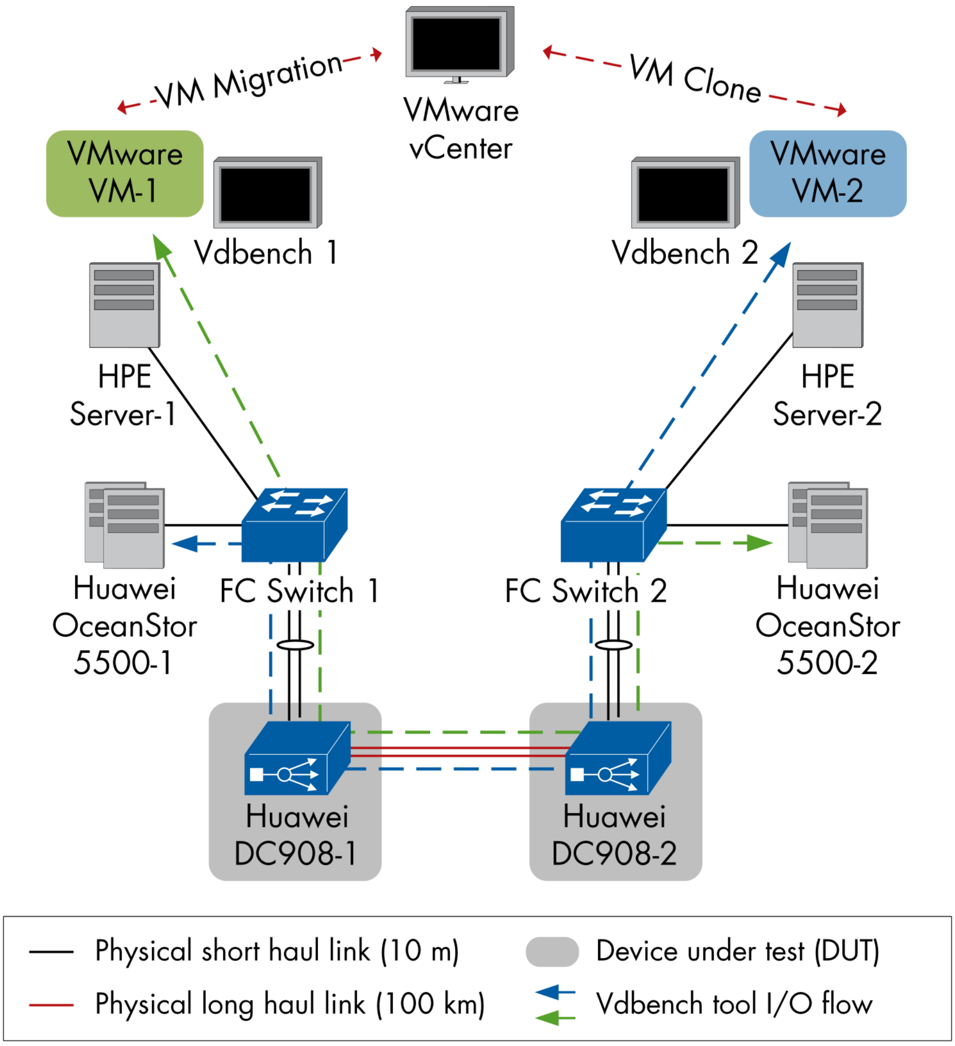


Figure 2: Logical Data Center Setup

Product Type	Product Name	Software Version
Devices Under Test WDM Equipment	Huawei OptiXtrans DC908	V100R019C10
	MD02A (line card facing FC switch)	
	OLL (line card facing WDM)	
	OPC (line card facing WDM)	
Physical Server	HPE DL380 Gen9	
Virtualization Platform	VMware vSphere	6.7
SAN Storage	Huawei OceanStor 5500 V5	V500R007C30
FC Switch	Brocade 6505	8.0.1
	Brocade G620	8.0.1
	Cisco MDS 9148S	6.2
	Cisco MDS 9132T	8.2

Table 3: Hardware and Software Components

Test Environment and Test Tools

VM clone and migration operations were conducted for the connectivity testing. The operations were managed by VMware vCenter. vCenter triggered the test VM provisioning on one of the data centers, including local storage. Once the test VM had been started, we created 200 GB large test files locally on the test VM. The test VM with these large size test files was cloned and/or migrated to a host and storage of the remote data center, connected through FC switches and Huawei DC908.

The test tool Vdbench used for the performance test and the long-run stability testing. The Vdbench was running in the test VM with the following configurations:

- Guest OS: CentOS 6.10, 64-bit
- Virtual disk: one TB Mapped Raw LUN
- Number of virtual disks: Five
- Number of virtual CPUs: Ten
- Guest memory: 200 GB
- Host Bus Adapter (HBA): one QLE2692 16Gb Fibre Channel to PCIe Adapter

The baseline traffic was generated with Vdbench that was running on one of the five Raw Device Mappings (RDM). The other four RDMs were used to generate the baseline traffic with the block size of 20 MB at maximum speed.

Test Results

Interoperability Testing

Firstly, we focused on the interoperability testing to verify hardware compatibility of the multiple optical transceivers between FC Switches and WDM systems.

FC-PI-6 Tests

Test Case	FC Link	VM Migration and VM Clone	
		Setup 1	Setup 4
		Brocade G620 pair	Cisco MDS 9132T pair
E-Port Disable/Enable Test	32G	Pass	Pass
	16G	Pass	Pass
	8G	Pass	Pass
Long Haul Network Failure	32G	Pass	Pass
	16G	Pass	Pass
	8G	Pass	Pass
ISL Trunking Test	32G	Pass	Pass
	16G	Pass	Pass
	8G	Pass	Pass
E-Port Cable Disconnect/ Reconnect Test	32G	Pass	Pass
	16G	Pass	Pass
	8G	Pass	Pass
Long Haul Network Failure with Protection	32G	Pass	Pass
	16G	Pass	Pass
	8G	Pass	Pass
Switch Reboot Test	32G	Pass	Pass
	16G	Pass	Pass
DUT Reset Test	32G	Pass	Pass
	16G	Pass	Pass
	10G	Pass	Pass
	8G	Pass	Pass
Extended Duration I/O Test	32G	Pass	Pass

Table 4: FC-PI-6 Tests

FC-PI-6 supported up to 32G link speed, and maintained backward compatibility to 16G and 8G link speeds. We selected 32G link speed for all operations. Additionally, we selected 8G and 16G link speeds for tests performing administrative operations in the optical domain across the WDM link.

The predefined FC-PI-6 optical transceivers connected in FC Switches and WDM systems, and then ISL established between FC Switches and WDM systems. We used VM migration and VM Clone to verify the interoperability with the VM operation traffic from VMware vCenter. We expected the successful ISL establishment, VM migration and VM Clone. The tests pass when all these expectations are achieved.

FC-PI-5 Tests with Hybrid Fiber Channel Switch Pair

For backwards compatibility in a hybrid Fibre Channel switch environment, as defined in setup 2 and 4 (see table 2 above), we selected 8G and 16G link speeds for the long haul link failure test and the line card removal test that challenge WDM in the optical domain.

The same processes and expectations used in this predefined FC-PI-5 tests to the above ones of FC-PI-6 tests.

FC-PI-3 Tests

For 10G ISL testing, we selected FC-PI-3 with the Brocade G620 pair. We verified the E_Port removal test, the long haul link failure test and the line card removal test.

The same processes and expectations used in this predefined FC-PI-3 tests to the above ones of FC-PI-6 tests.

Test Case	FC Link	VM Migration and VM Clone	
		Setup 2	Setup 5
		Brocade 6506 - Brocade G620	Cisco MDS 9148S - Cisco MDS 9132T
Long Haul Network Failure	16G	PASS	PASS
	8G	PASS	PASS
DUT Reset Test	16G	PASS	PASS
	8G	PASS	PASS

Table 5: FC-PI-5 Tests with Hybrid Fibre Channel Switch Pair

Test Case	FC Link	VM Migration and VM Clone
		Setup 3
		Brocade G620 pair
ISL Trunking Test	16G	PASS
E_Port Fiber Disconnect/Reconnect Test	8G	PASS
Long Haul Network Failure with Protection	16G	PASS
DUT Reset Test	8G	PASS

Table 6: FC-PI-3 Tests

Capability Testing

Secondly, we conducted functional capability tests. To verify the long-distance forwarding of the WDM between the data centers, we concentrated on administrative use cases to allow operators to perform maintenance activities. We simulated customer-facing operations in the optical domain, such as changes in the E_Port configuration of the FC switches, disconnection of fibers from the FC switches facing WDM system, and even the FC switch reboot while the WDM was forwarding. We carried out physical operations on the WDM system, such as removal and reinsertion of the line card. In this domain, we looked at the link protection test by the ISL trunking to protect the WDM against link failure.

E_Port Disable/Enable Test

We verified the interoperability of the WDM solution with four FC switches at each with the defined speed. The focus of this test is that when the state of the E_Port changes, the WDM system remains transparent and will not affect the state of the E_Port. As specified in the Fibre Channel interconnect protocol (vendor-proprietary, supported by Brocade and Cisco), an ISL implements and terminates at an E_Port on each end. In the data center under test, the WDM link between FC switches implements the ISL link between the FC switches and is connected to the optical domain with the E_Port. A F_Port on the FC switch is used to connect with a storage system or a physical server.

We performed a baseline test to verify the throughput of the interface along with FC switches through the WDM system. The baseline test was executed in setup 1 and 4. The performance results are shown below.

While running the baseline traffic, we disabled the two E_Ports on the both FC switches in two data centers. After observing the decreased traffic rate in storage traffic and error log from the FC switch reporting the port went down to confirm that the configuration took effect, we performed a VM migration which failed as expected. Then, we activated the port in the configuration and observed that the VM migration successfully completed indicating the restored status of the E_Port. The performance of the maximum throughput is shown below.

FC Link	Setup 1		Setup 4	
	Expected Throughput (GB/s)[5]	Measured Throughput, per direc-	Expected Throughput (GB/s)	Measured Throughput, per direc-
32G	3.104	1.539[6]	3.104	1.539[6]
16G	1.552	1.539	1.552	1.536
8G	0.776	0.769	0.76	0.745

Table 7: Achieved Speed Running Vdbench

[5] The expected throughput is based on layer 2 payload, we used a ration of 97% at link speed. The ratio is based on the formular: 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.

[6] The sent traffic was equal to 1.6 GB/s as thjs is the maximum provided by the Huawei OceanStor 5500 V5 equipped with 20 serial attached SCSI (SAS) hard disk drive (HDD) disks.

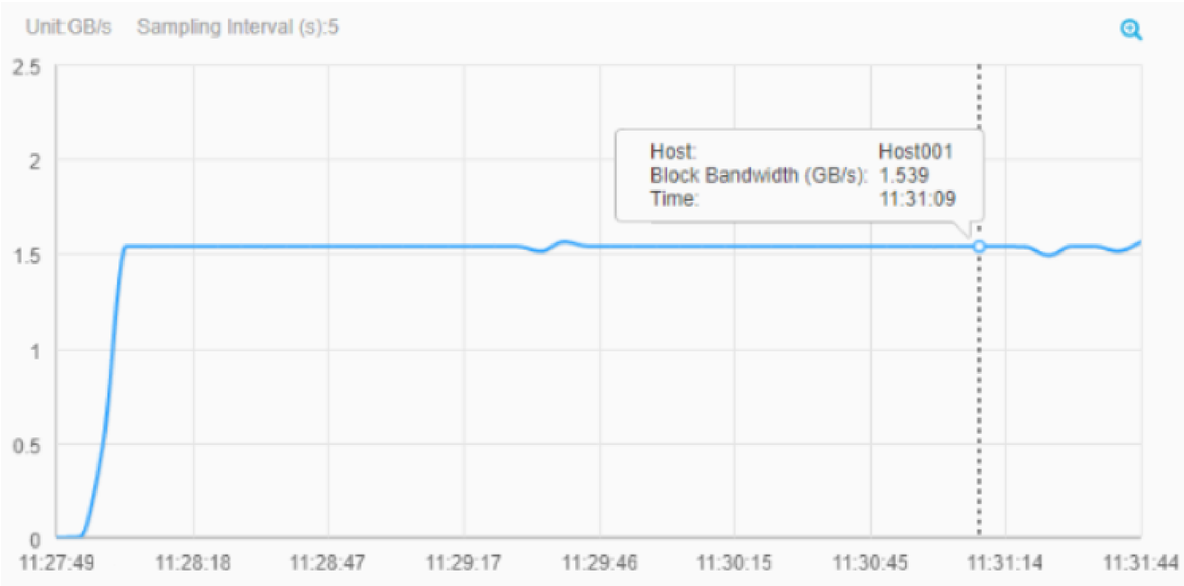


Figure 3: Maximum Throughput at 1.539 GB/s

E_Port Fiber Disconnect/Reconnect Test

In this test, we verified that the WDM system remains transparent when the E_Port fibers are disconnected and reconnected from the FC switch. To achieve repeatability we used the same baseline traffic from the previous test case. The baseline test was executed in setup 1, 4 and 5, the performance result is as follows.

We unplugged all fibers at the E_Ports on the current FC switch. After observing the decrease in traffic in the storage system and the error log of the FC switch reporting that the port had an expected failure, the VM operation failed as expected, so we plugged in the fiber. As expected, the VM returned to the operating state and the VM migration and cloning continued.

ISL Trunking Test

We verified the link protection between the WDM system and the FC switch. The traffic continued on the remaining link when one link failed. While running baseline traffic, we removed one of two running fibers from ISL trunk and observed that the VM operation was continued on the remaining link as expected. The result of the switch over time is shown below.

After we reconnected the link to the ISL trunk we did not observe any impact on the baseline traffic.

FC Link	Setup 1		Setup 4		Setup 3	
	Expected Throughput (GB/s)	Measured Throughput, per direction (GB/s)	Expected Throughput (GB/s)[5]	Measured Throughput, per direction (GB/s)	Expected Throughput (GB/s)[5]	Measured Throughput, per direction (GB/s)
32G	3.104	1.539[7]	3.104	1.539[7]	Not applicable	
16G	1.552	1.536	1.552	1.521		
10G	Not applicable		Not applicable		1.164	1.156

Table 8: Achieved Speed Running Vdbench

[7] The sent traffic was equal to 1.6 GB/s as maximum provided by the Huawei OceanStor 5500 V5 equipped with 20 serial attached SCSI (SAS) hard disk drive (HDD) disks.

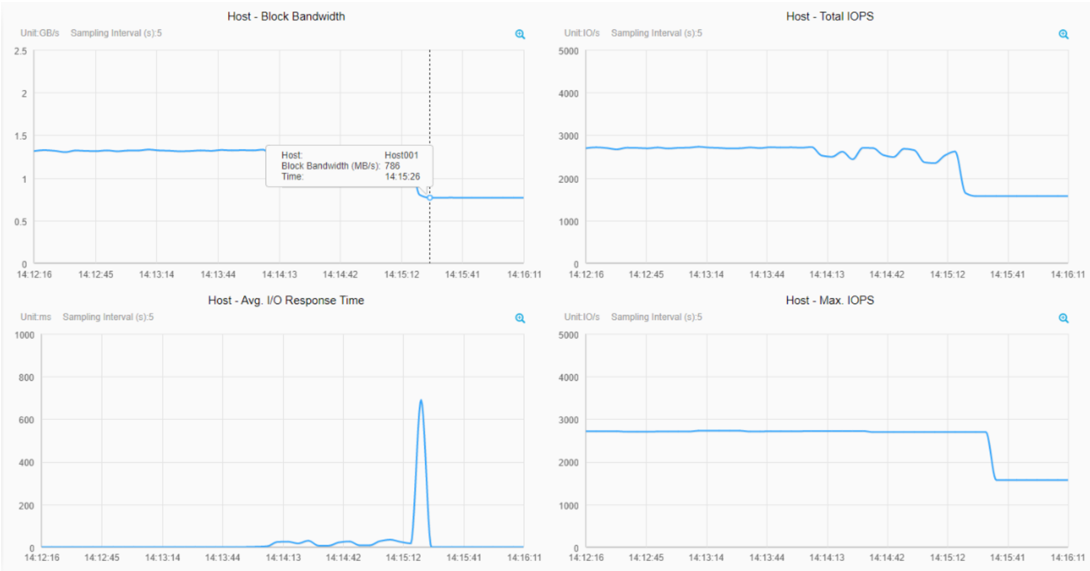


Figure 4: Switch Over Time in Storage GUI Graph

DUT Reset Test

We verified that the DUT forwards storage data after the line card has been re-inserted.

While running baseline traffic, we removed the MD02A line card from the Huawei DC908. We observed that the VM operation failed during the removal of the MD02A line card. We reinserted the line card into the DUT and observed a successful VM migration.

FC Link	Setup 2		Setup 5	
	Expected Throughput (GB/s)[5]	Measured Throughput, per direction	Expected Throughput (GB/s)	Measured Throughput, per direction
16G	1.552	1.454[8]	1.552	1.414[8]
8G	0.776	0.745	0.776	0.727

Table 9: Achieved Speed Running Vdbench

[8] The traffic rate results proved our speculation that 100 km long haul is not recommended for setup 2 and 5 (the required buffer size not supported on the Brocade 6505 and the Cisco MDS 9148S).

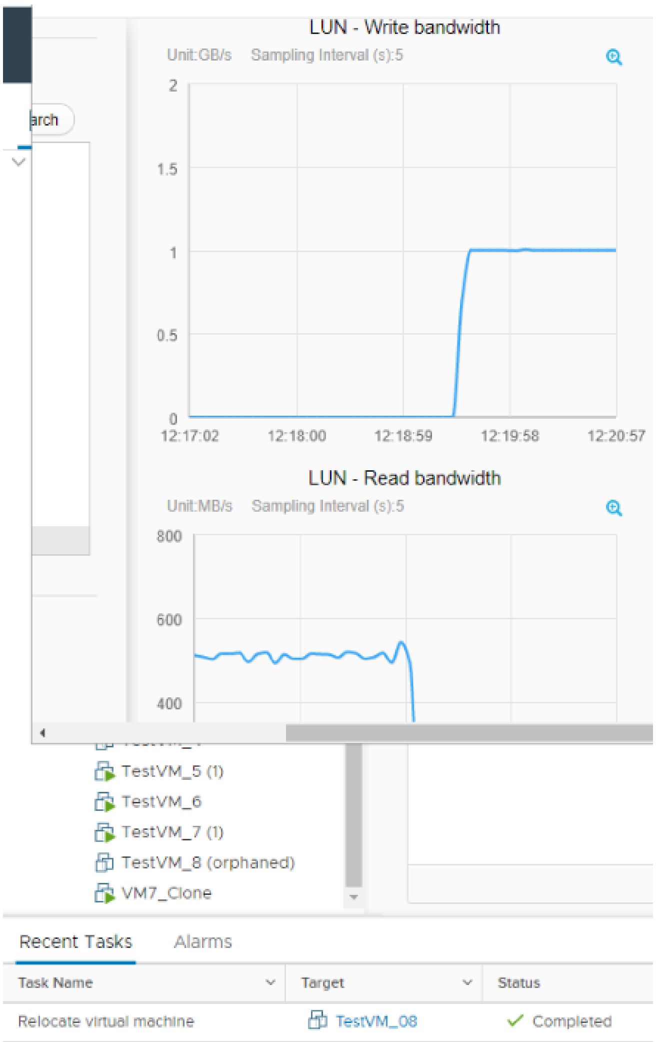


Figure 5: Normal VM Operation after the Resetting

Long Haul Network Failure

We validated that the WDM system continues to forward storage data after long haul links failure. We used the same baseline traffic from the previous test case to achieve repeatability. While running baseline traffic, we removed both long haul fibers from the Huawei DC908 and observed a failed VM operation and I/O operations of storage. We reconnected the long haul fiber and observed a successful VM clone.

The baseline test was executed in all four setups. The performance result of setup 2 and 5 is shown below. The performance result of setup 1 and 4 can be found above in Table 7.

Long Haul Network Failure with Protection

We verified that the WDM link protection and traffic switches over to the backup link when the primary link fails.

While running baseline traffic, we removed one of the two long haul fibers from the Huawei DC908 and observed a shortly impacted VM operation and I/O operations of storage as expected. We only kept one long haul fiber and observed successful VM clone.

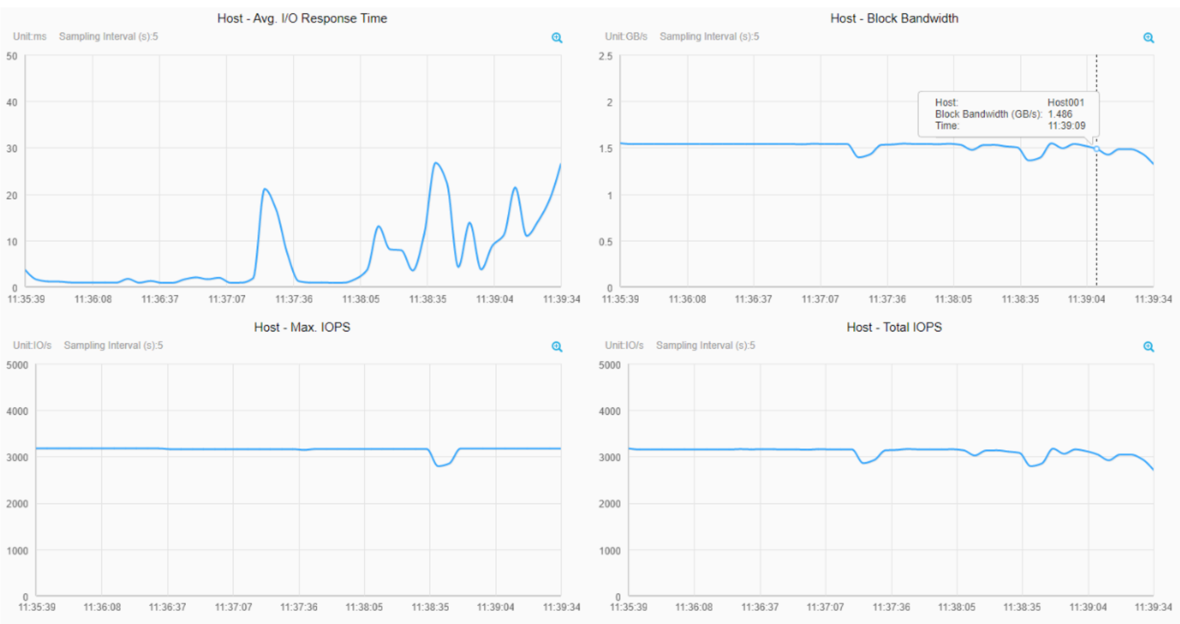


Figure 6: Link Fail Over Time in Storage GUI Graph

Switch Reboot Test

We verified that the WDM system forwards storage data after the switch connected to the MD02A line card was rebooted.

While running benchmarking traffic, we restarted the FC switch connected to one of the Huawei DC908 devices. As repeatedly observed, the VM operation failed and the I/O operation throughput decreased. With the switch up and running again, we repeated the baseline throughput test and observed that the VM migration was completed successfully.

Stability Testing

To verify the stability of the WDM between data centers, we soaked it with the continuous baseline traffic for 24 hours.

Extended Duration I/O Test

We verified the reliability of the WDM system in terms of performance consistency under the conditions of long period stress load.

The test tool Vdbench triggered the baseline traffic bidirectionally for 24 hours. During that time, we monitored the system log of both hardware and software. We confirm that the system is able to transfer the data by a consistent rate for a duration of 24 hours. Additionally, we didn't observe any software crashes or hardware failure during the test duration.

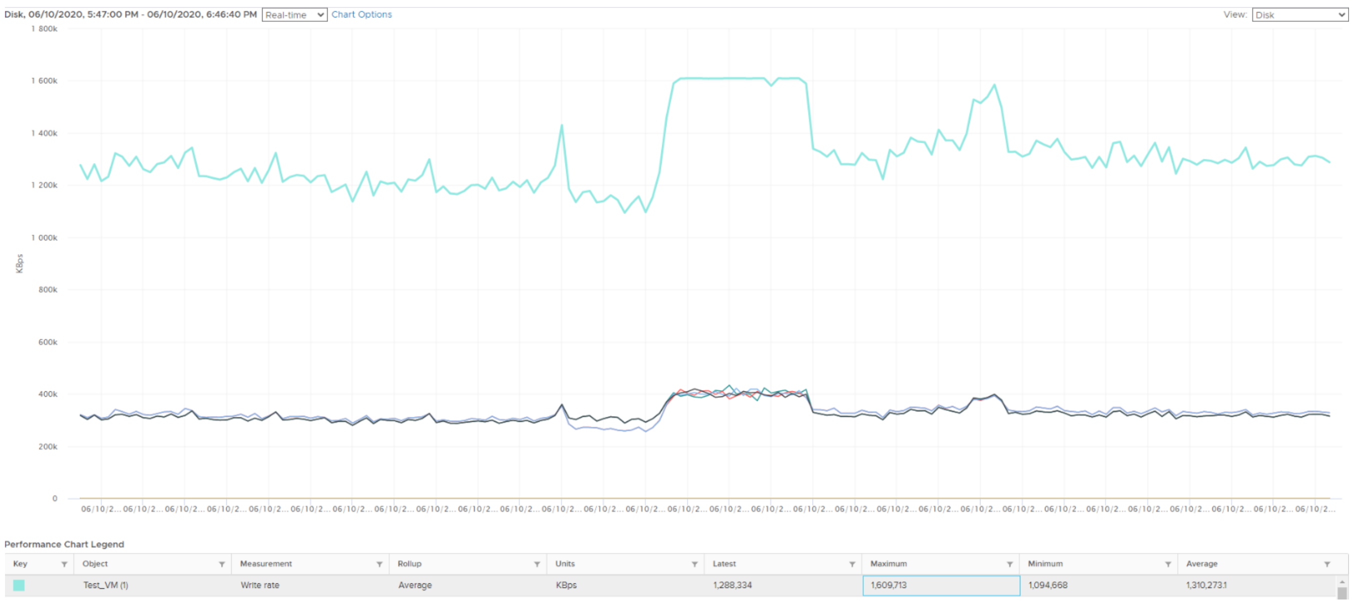


Figure 7: 24 Hours Load Test Result

2021

The following sections describe the test results of the software version V100R021C10.

Testbed Description

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 Ocean-Stor 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.

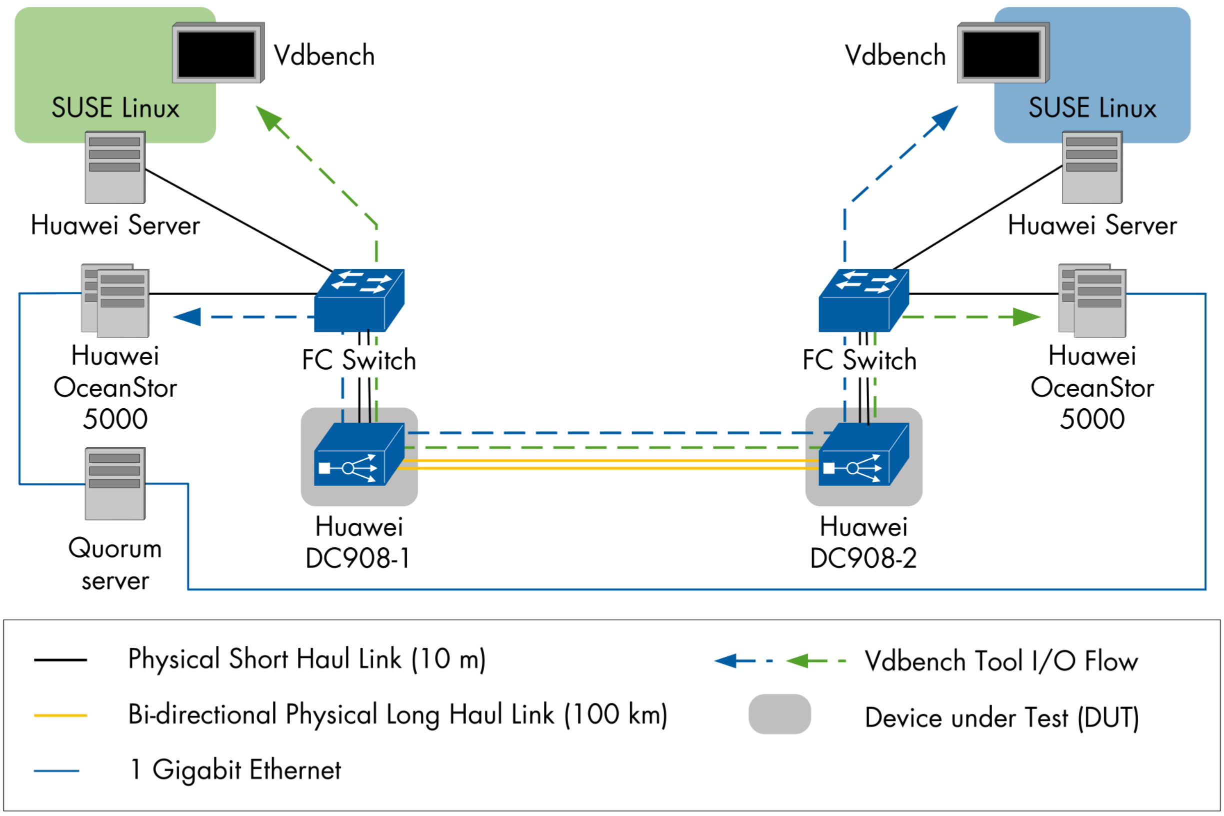


Figure 8: Logical Test Topology

Product Type	Product Name	Software Version
Devices Under Test WDM Equipment	Huawei OptiXtrans DC908	V100R019C10
	MD02A (line card facing FC switch)	
	OLL (line card facing WDM)	
	OPC (line card facing WDM)	
	SCC Panel	
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[9]	V300R002C10
Quorum server	Quorum server	V300R002C10
FC Switch	Brocade 6505	FOS v8.2.1a
	Brocade G620	FOS v9.0.1c
	Brocade 300	FOS v7.4.1f

Table 10: Hardware and Software Components, Test Environment

[9] 24 Non-Volatile Memory Express (NVMe) solid-state disk (SSD) disks that provided up to 5.8 GB/s (Gigabytes per second) Input/Output

Capacity Testing

We measured the maximum FC-interface forwarding of the DUT using 1, 2, 4, 8, 10, 16, and 32 Gbytes/s block traffic generated respectively between emulated data centers, which are 100KM away from each other.

The following tables show the throughput. We performed a capacity test for the MD02A line card under test, with all three setups, respectively. Each traffic stream carried bi-directional traffic.

Setup	Speed Type	Expected Throughput (MB/s)[10]	Measured Throughput, per direction (MB/s)	Verdict
1. Brocade 300	FC100	100	96	Pass
	FC200	200	194	Pass
	FC400	400	388	Pass
2. Brocade G620 pair	FC800	800	781	Pass
	FC1200	1200	1168-1170	
	FC1600	1600	1556-1562	
	FC3200	3200	3103-3105	Pass
3. Brocade G620/6505 pair	FC1600	1600	6505: 1530 G620: 1560	Pass

Table 11: MD02A Interface Throughput

10 The expected throughput is based on layer 2 payload, we used a ration of 97% at link speed. The ratio is based on the formular: 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 created a new test setup with the Viavi MTS5800-100G tester to measure the latency value of the WDM system.

In this setup, we connected both DUTs back to back with the measurement tool (see Figure 7), which provides latency value in the precision of microseconds. The latency indicated the whole system latency, including MD02A, EMR8, OPC, and SCC.

We removed all data center devices from the test bed and remained only two 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, respectively. 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 connected them to WDM devices.

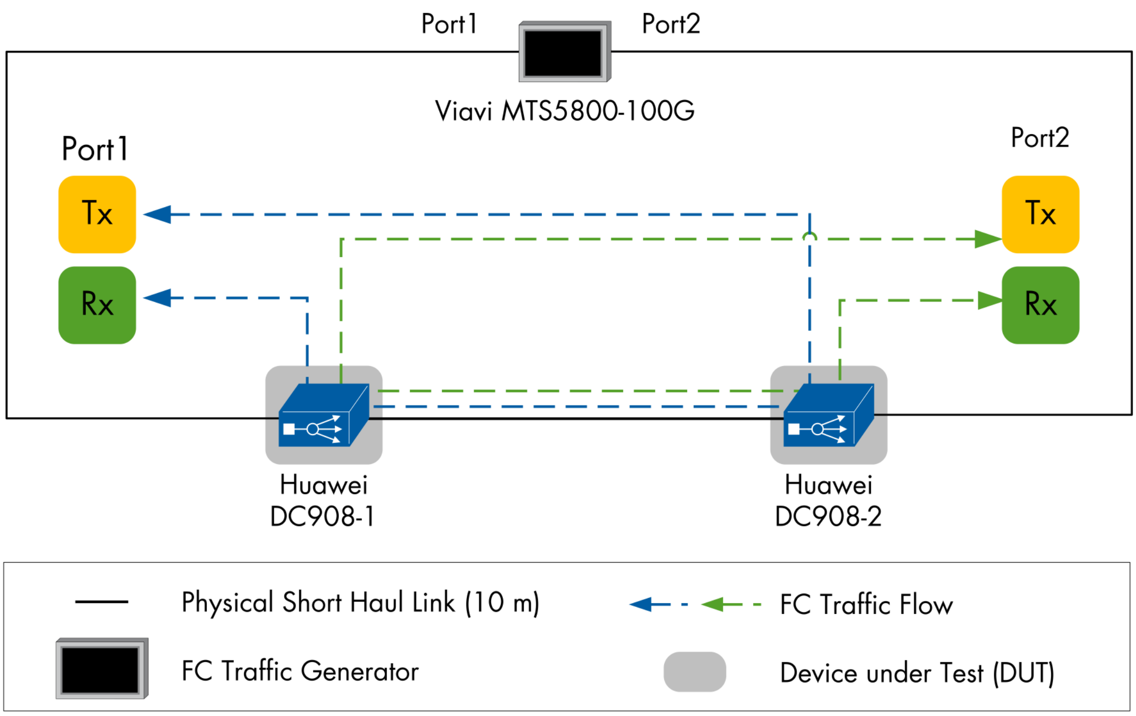
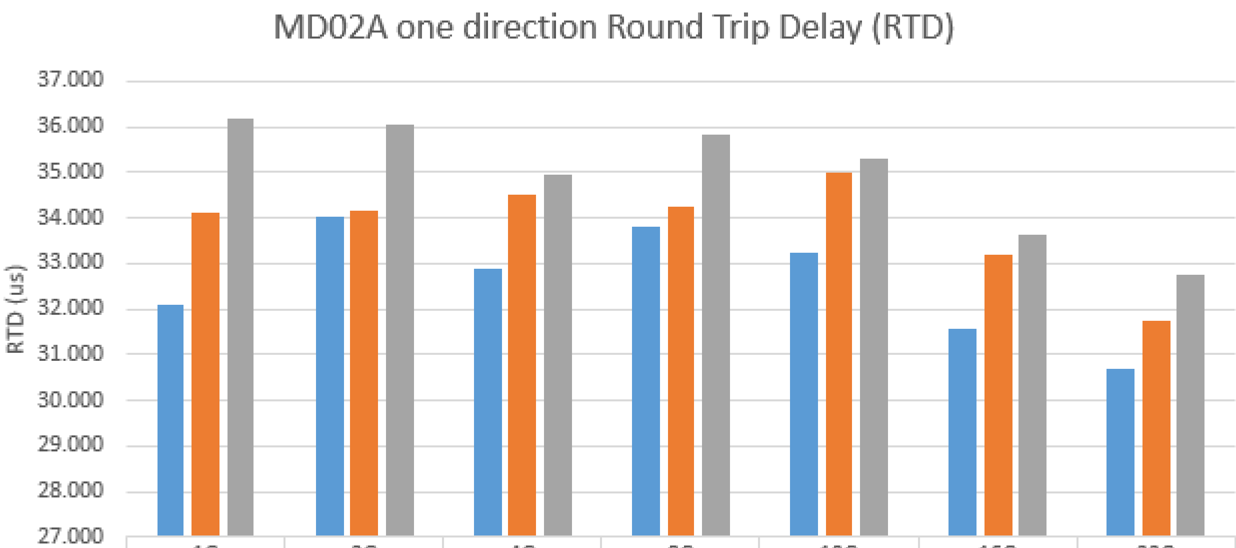


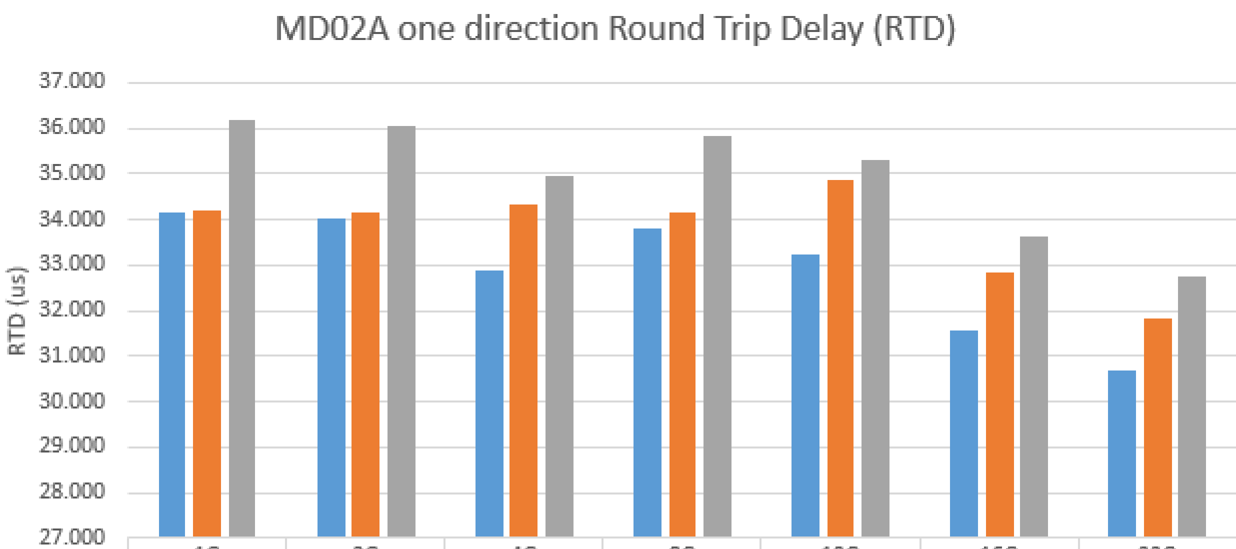
Figure 9: Latency Test Topology

The figure below shows the FC traffic running from DUT1 to DUT2 and the reverse direction. The maximum latency was identical in both directions, as expected.



	1G	2G	4G	8G	10G	16G	32G
Min.	32.102	34.013	32.885	33.804	33.252	31.581	30.701
Avg.	34.120	34.165	34.520	34.245	35.008	33.181	31.755
Max.	36.198	36.061	34.932	35.852	35.300	33.629	32.749

Min. Avg. Max.



	1G	2G	4G	8G	10G	16G	32G
Min.	34.150	34.013	32.885	33.804	33.252	31.581	30.701
Avg.	34.193	34.161	34.332	34.150	34.882	32.855	31.843
Max.	36.198	36.061	34.932	35.852	35.300	33.629	32.749

Min. Avg. Max.

Figure 10: Latency Results of DC908 + MD02Asystem (DUT1 to DUT2, DUT2 to DUT1)

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 the local side of the DCI initiated I/O operations, and the target was the remote 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 six 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.

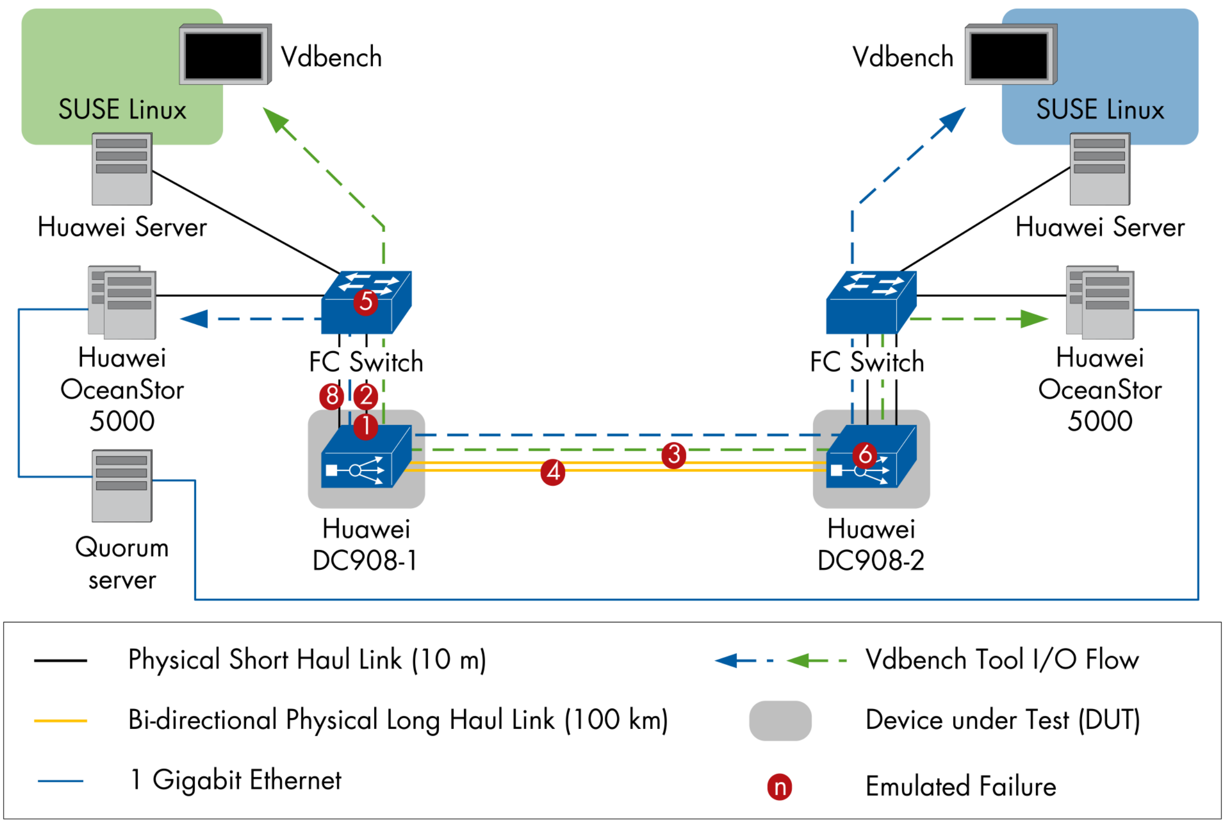


Figure 11[11]: Emulated Failure in the Test Bed

[11] The numbers in figure 11 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 with the host. The secondary storage stopped receiving any I/O operation from the remote data center or the local data center. The former case was that these four types of failure could interrupt the whole DCI connection from different hardware locations (see Figure 12), causing it to stop transmitting data between both data centers. The latter phenomenon was due to the design mechanism.

Huawei OceanStor 5000 V3 implements a technology called "HyperMetro" for synchronizing the states between preferred storage and secondary 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. Once the quiescing time (300 s) passed during a link recovery, we observed that the traffic switched back to baseline status between DC1 and DC2. We did not observe any impact as expected.

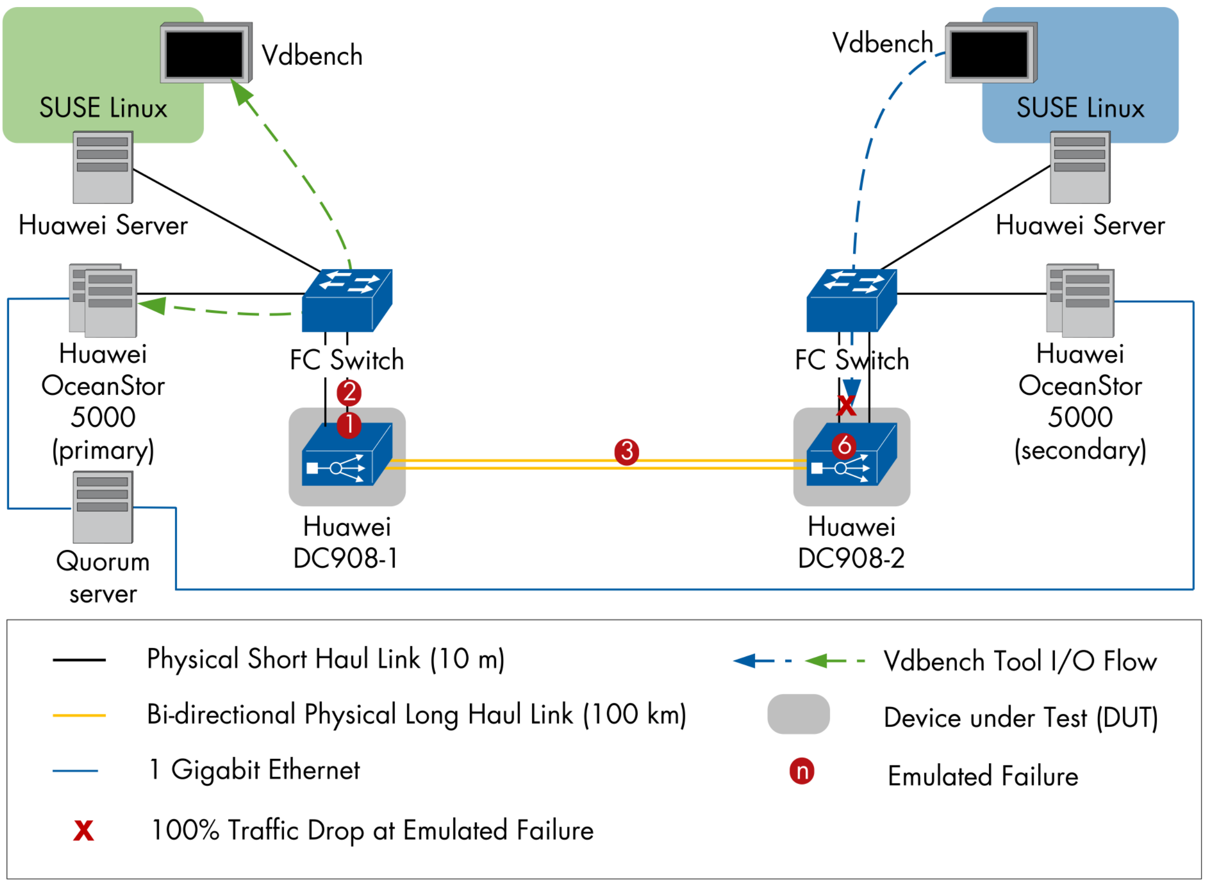


Figure 12: Traffic Status after the Switch Over for Failure Types 1, 2, 3 and 6

No. in Figure 12	Test Case	Action	Service Interruption		Verdict
			Expected[12]	Observed (for primary storage)[13]	
1	E-Port Disable/Enable Test	Disable	10s	7s including 5s drop to 0MB/s	Pass
		Enable	No impact	No impact	Pass
2	E-Port Cable Disconnect/Reconnect Test	Disconnect	10s	7s including 5s drop to 0MB/s	Pass
		Reconnect	No impact	No impact	Pass
3	Long Haul Network Failure without Redundancy	Disconnect	10s	7s including 5s drop to 0MB/s	Pass
		Reconnect	No impact	No impact	Pass
6	DUT Reset Test	Disconnect	10s	7s including 5s drop to 0MB/s	Pass
		Reconnect	No impact	No impact	Pass

Table 12: Out of Service Overview (Active-Active Storage Cluster Protection)

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 continued to perform I/O operations. In the data center where the FC switch rebooted, its connected storage stopped receiving any data from the FC switch as long as the reboot took place. This behavior is expected. After that, the FC switch completed the reboot. 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.

As we have three pairs of FC switch, which is Brocade 300 / Brocade 300, Brocade G620 / Brocade G620, and Brocade G620 / Brocade 6505. They have different software versions and hardware modules. The behavior after the reboot was different. The service of rebooted Brocade 300 had totally stopped on the reboot side until the reboot procedure was done and the quiescing time had passed. The service of rebooted Brocade G620 had not fully stopped on the reboot side. However, the FC traffic did stop for around 40-50 seconds until it recovered and read the non-rebooted side storage. Everything was recovered after the quiescing time had passed.

The interesting part is Brocade G620 and 6505 pair. The service interruption depended on which device we rebooted. If we reboot G620, the behavior is the same as G620's. If we reboot 6505, the behavior is the same as 300's.

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

[13] We observed 5 s complete traffic drop to 0 MB/s during the switch over. The Huawei team configured the quorum server with 5 s 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 6 s complete traffic drop to 0 MB/s included the time from discovery to switchover at the quorum server.

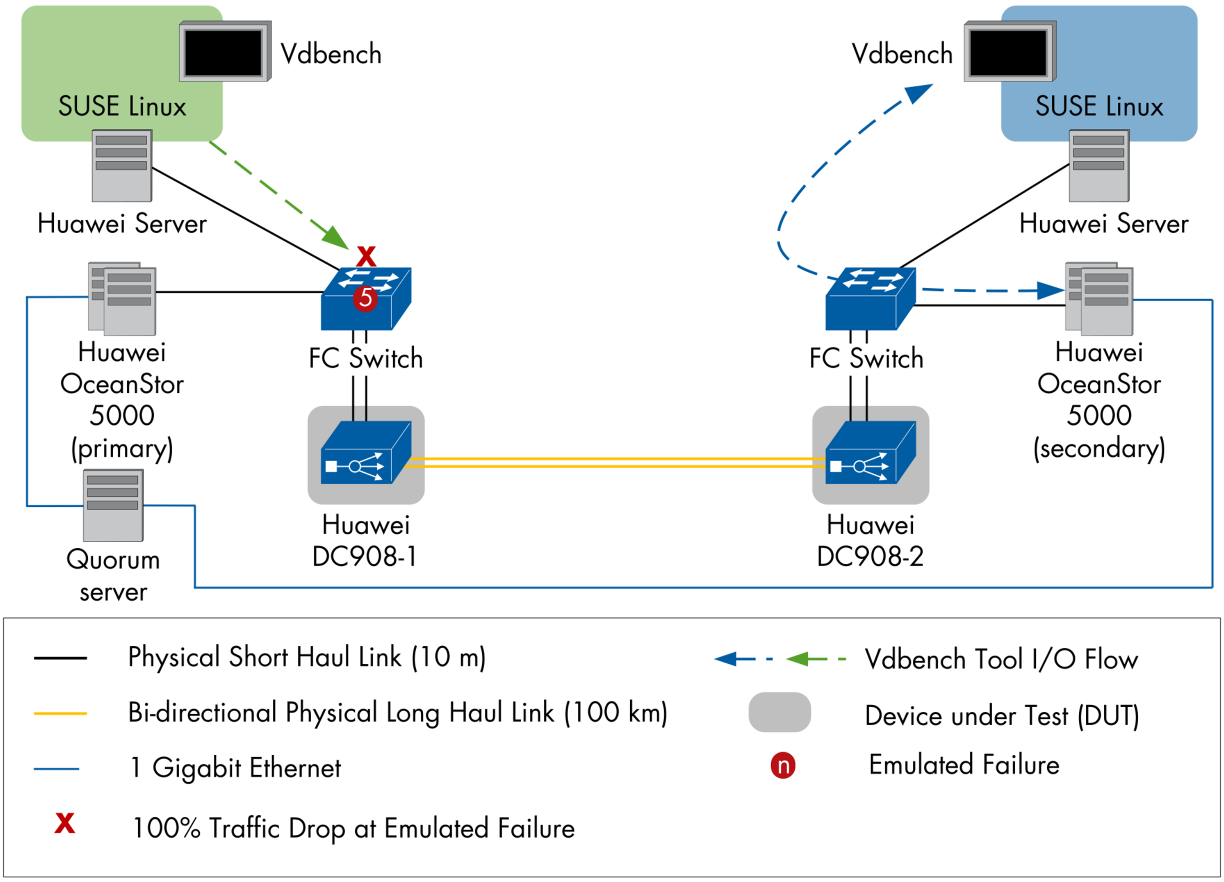


Figure 13: Traffic Status during FC Switch's Reboot

No. in Figure 13	Test Case	Action	Service Interruption		Verdict
			Expected	Observed (for primary storage)[15]	
5	Switch Reboot Test	Shut down	15s[14]	10s including 8s drop to 0MB/s	Pass
		Turn on	No impact	No impact	Pass

Table 13: Out of Service Overview

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

[15] The 8 s 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 of the DUT that switches 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. Table 11 shows the measured switch over 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 plug the fiber back to its original port, and we expect the traffic to increase by 100%.

No. in Figure	Test Case	Action	Service Interruption		Verdict
			Expected	Observed	
4	Long Haul Network Failure with Redundancy	Disconnect	4s[16]	4s including 2s drop to 0MB/s (for both primary and secondary storages)	Pass
		Reconnect	No impact	No impact	Pass
8	ISL Trunking	Disconnect	Drop 50%	Drop 50%	Pass
		Reconnect	Increase 100%	Increase 100%	Pass

Table 14: Out of Service Overview—Setup Brocade G620 pair

FC switch pairs	Switch Over Time (s)
Brocade 300/Brocade 300	A maximum of 4s service interruption (including a maximum of 2s drop to 0MB/s)
Brocade G620/Brocade G620	A maximum of 3s service interruption (including a maximum of 1s drop to 0MB/s)
Brocade G620/Brocade 6505	A maximum of 3s service interruption (including a maximum of 1s drop to 0MB/s)

Table 15: Switch Over Time of MD02A

16 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 500 ms (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.

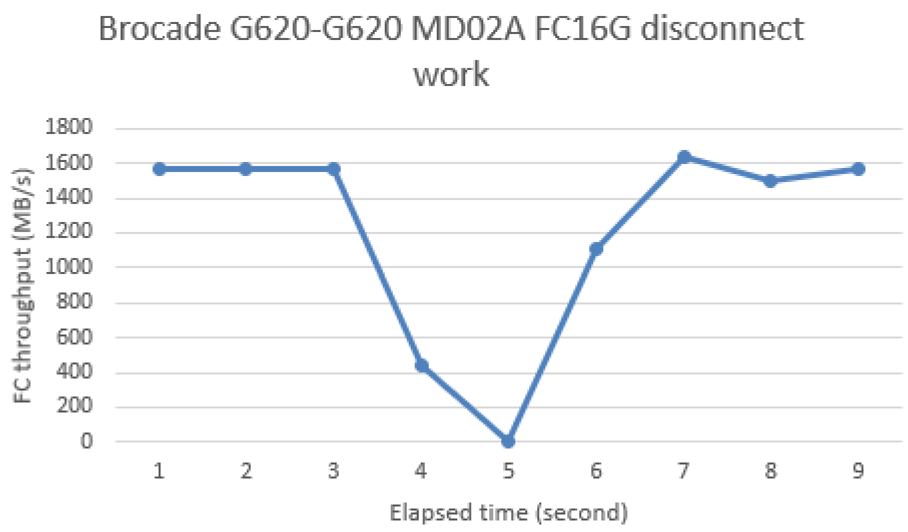


Figure 14: WDM Link Switch Over with Brocade G620

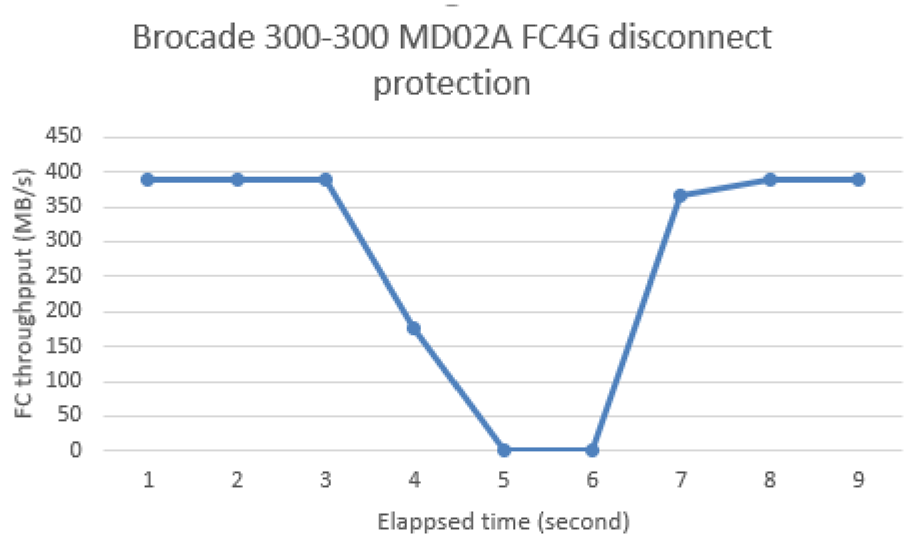


Figure 15: WDM Link Switch Over with Brocade 300

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 test setup. We used the maximum load of the baseline traffic for each selected interface as measured in capacity tests and formed 24 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
MD02A	1G	2G	1G	4G	2G	4G

Table 16: Brocade 300 Pair

Board Under Test/ Test Scenario	E-port Interface used in Test Case		
	1	4	6
MD02A	16G	16G	16G

Table 17: Brocade G620 - 6505 Pair

Board Under Test/Test Scenario	E-port Interface used in Test Case							
	1	2	3	4	5	6	7	8
MD02A	8G, 16G	10G, 32G	8G, 32G	10G, 16G	8G, 16G	32G	16G	16G, 32G

Table 18: Brocade G620 Pair

Soak 24 Hours Test

We verified the WDM system's reliability in performance consistency under long-period stress load conditions. The Huawei team configured three chassis 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 consistently 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.

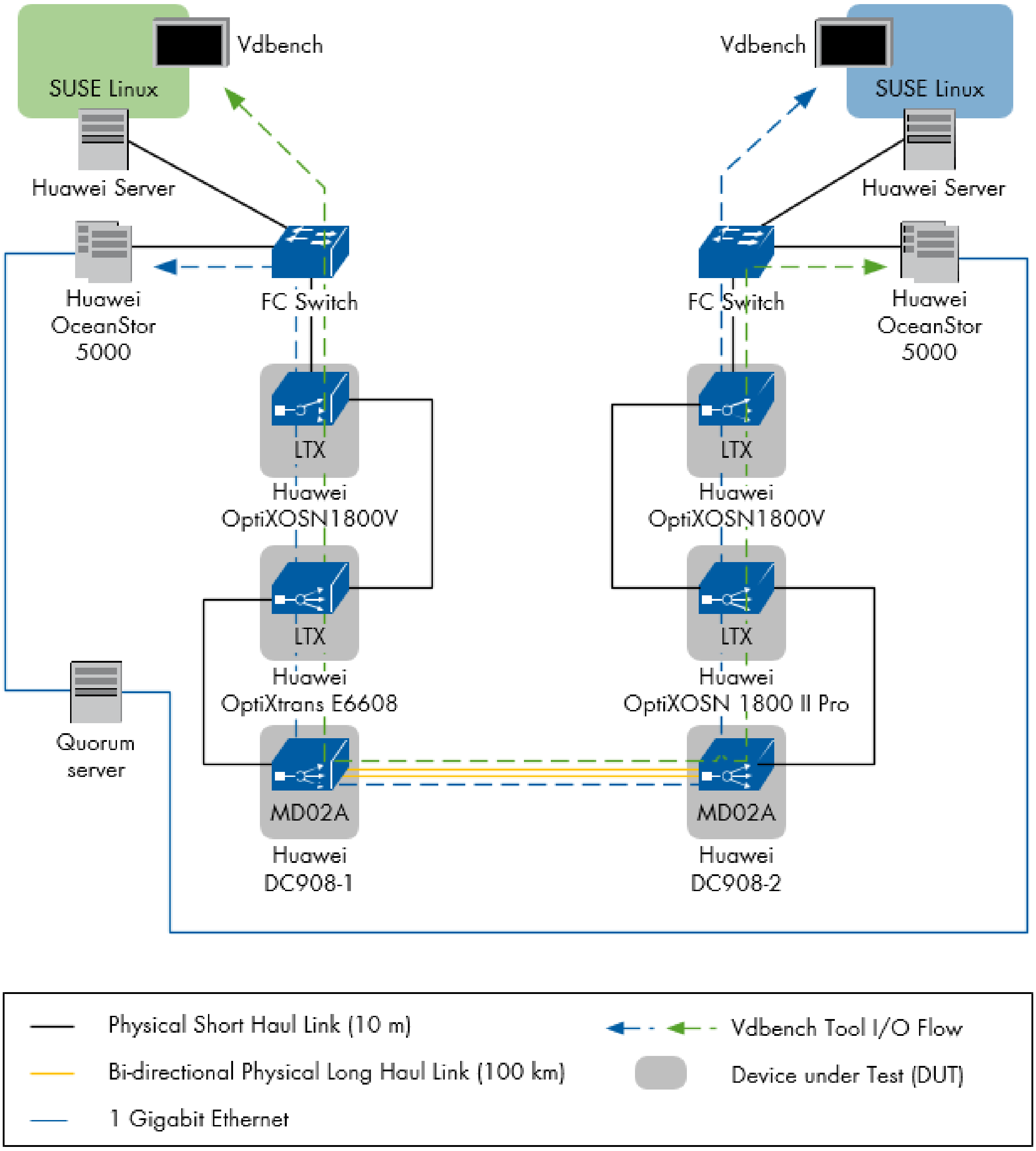


Figure 16: Topology of 24-hour Throughput at 16G

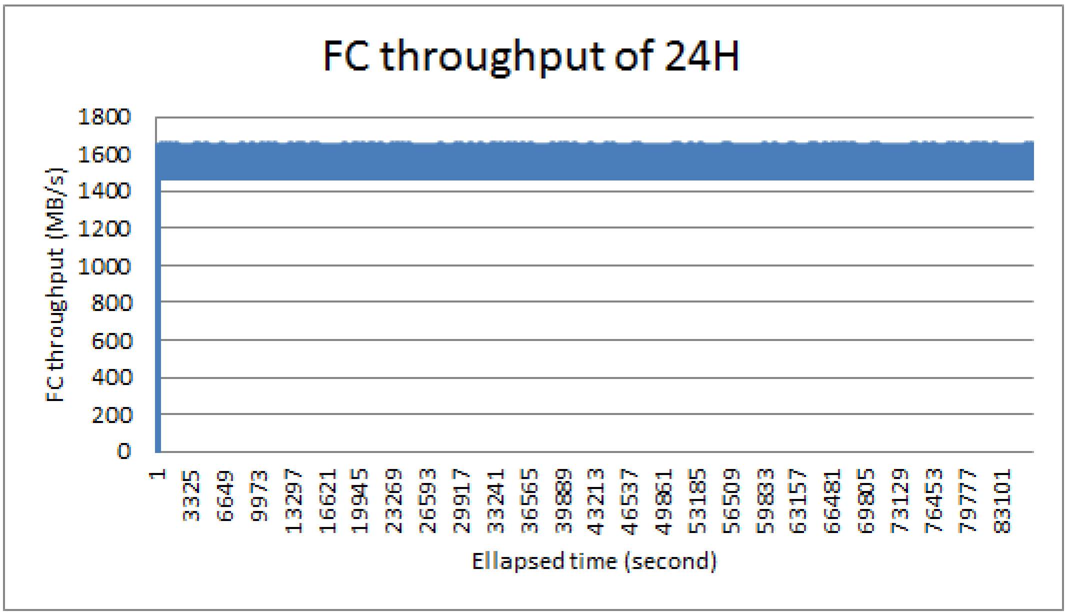


Figure 17: 24-hour Throughput at 16G

Conclusion

We verified interoperability of Huawei DC908 and third-party Fibre Channel switches from Brocade (6505 and G620) and Cisco (MDS 9148S and MDS 9132T) in 2020 and Brocade (300 / 300, G620 / G620, G620 / 6505) in 2021. Multiple optical transceiver functions were certified in the E_port between these FC switches and the Huawei DC908, including FC-PI-6, FC-PI-5, and FC-PI-3.

EANTC validated forwarding speeds in FC 8G, 10G, 16G, and 32G scenarios in 2020 (software version V100R019C10). With the latest software version in 2021 (V100R021C10), MD02A supports 1G, 2G, 4G as well. When forwarding traffic at any of these standardized speeds, the Huawei DC908 did not exhibit any speed impact. The extended 24 hours soak testing confirmed stable operations of DC908 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 MD02A 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 OptiXtrans DC908 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]