

Fibre Channel over Ethernet – A Practical Alternative

Executive Summary

Based on comprehensive testing performed by Cisco and NetApp, this white paper explores whether Fibre Channel over Ethernet (FCoE) is a viable and practical alternative to current Storage Area Networking (SAN) solutions such as Fibre Channel (FC) and Internet Small Computer System Interface (iSCSI). The testing clearly demonstrated that FCoE has numerous advantages, in that it allows Local Area network (LAN) and SAN communications to share a common infrastructure, eliminating the need for independent cabling and parallel networks, while preserving the tried and tested SAN provisioning and management methodologies used by storage administrators today. The testing also showed that FCoE performed at least as well as FC and consumed less server resources than either iSCSI or NFS (the current alternative solution for consolidating LAN and SAN traffic). Because of the potential capital and operational cost savings associated with a unified networking solution for SAN and LAN communications and the relative ease of migration to FCoE, we would recommend that data center designers and operations teams consider the use of FCoE, in conjunction with 10 Gigabit Ethernet, for their future Storage Area Networking requirements. Please note that any performance figures quoted in this document are for implementation comparison purposes only, (based on the single host test environment used) and not a measure of the maximum performance and throughput possible in a larger scale implementation.

What You Will Learn

Many organizations are carefully watching developments of FCoE, a new storage networking technology recently standardized by the International Committee for Information Technology Standards (INCITS) T.11 in the FC-BB-5 specification. Data center managers and designers are extremely interested in the technology, as it offers the potential for consolidating both LAN and SAN traffic onto a single, unified network (or fabric, to use SAN terminology). This development could drastically reduce the capital outlay associated with building data center network infrastructure, and could deliver tangible operational expenditure savings associated with running and operating data center connectivity.

To determine whether FCoE is a viable technology, the operational characteristics also need to be considered. FCoE needs to perform as well as FC, and it needs to be able to be managed in the same way. Additionally, performance metrics must be as good, or better, than those of FC.

The Cisco Enhanced Customer Aligned Testing Services (ECATS) team, in conjunction with NetApp, has recently undertaken a series of tests and assessments designed to determine whether FCoE, as it exists today, provides a practical and realistic alternative to current storage networking technologies such as FC and iSCSI, in terms of deployment, high availability, operating procedures, performance and diagnostics.

This white paper provides an overview of these tests, and presents the conclusions derived.

Introduction

Today, the vast majority of data centers have multiple, independent networks. Each of these networks is designed to deliver specific, optimized services to the computing and storage resources needed to run an organization. LANs based on Ethernet are commonly used to provide user access to application and file servers, while FC is the networking technology of choice for SANs that connect servers to disk arrays and tape libraries. Other networking technologies (e.g., InfiniBand) may also be present, dedicated to providing high bandwidth/low latency connections between servers. These can be used for cluster applications such as Oracle Real Application Clusters (RAC) or high performance computing (HPC) systems used in the scientific and financial sectors.

However, a recent trend has emerged whereby more and more services are being migrated to Ethernet technology. This trend has been driven by the desire to simplify data center infrastructures by eliminating the need for separate cabling and switching technology for different connectivity requirements of a server or storage system. For example, iSCSI (which allows SCSI mode disk access to run over Ethernet LANs using the TCP/IP protocol) is now used by many organizations to deliver block level storage array access for smaller servers, while 10 Gigabit Ethernet (GE) is seen as a viable alternative to InfiniBand for many clustering solutions. This transition minimizes complexity and thus reduces both the capital and operational expenditure associated with building and running a data center.

Unfortunately, especially with regard to SAN access, iSCSI is not considered a viable alternative to FC SAN fabrics by many organizations for a number of reasons. These include:

- Concerns over server performance being impacted by the need to use TCP/IP for storage access and flow control
- Lack of support for SAN boot over iSCSI
- Different operational administration practices and naming conventions between iSCSI and FC

Therefore, there is now a clear requirement emerging for a technology that allows SAN traffic to co-exist on the same network as LAN traffic, while simultaneously addressing all of the points listed above. This has resulted in the development of FCoE, a converged networking technology, by the INCITS T.11 organization (responsible for specifying FC SAN standards). FCoE is managed, configured and operates in the same manner as FC, yet runs on Ethernet networks. It enables the SAN traffic to co-exist on the same physical cable and switch infrastructure as the LAN traffic. In parallel to this effort, the Institute of Electrical and Electronics Engineers (IEEE, the organization responsible for LAN standardization) has been developing extensions to the Ethernet protocol to allow better differentiation of common traffic types within the data center and deliver different behavior based on the traffic type. For example, these extensions (collectively known as "Data Center Bridging") allow FCoE traffic to receive reserved bandwidth and "lossless" service (i.e., no traffic discarded) from Ethernet devices that are virtually identical to that provided by FC switches. Traditional LAN traffic can continue to experience traditional LAN behavior, where error detection and recovery is handled by upper layer protocols such as TCP/IP.

However, before converged networks utilizing FCoE will be accepted as viable alternatives to current deployments based on separate Ethernet LANs and FC SANs, data center designers and operations teams need to see that FCoE has the same operational characteristics as FC, and has performance metrics that are similar or better than current FC solutions. In order to assess and demonstrate that FCoE is in fact directly equivalent both in terms of operation and performance to current FC solutions, Cisco, in conjunction with NetApp, has carried out a detailed set of tests

analyzing and comparing FC and FCoE. These tests looked at normal operational activities associated with provisioning and managing a storage area network, and compared the respective performance of the two solutions. To provide a valid comparison of all current solutions, the tests also looked at both iSCSI and Network File System (NFS) based solutions.

Testing Topology and Configuration

We recognize that an immediate, full-scale migration to an FCoE environment may not be desirable for many companies, especially for those who have heavily invested in non-FCoE-based solutions. Therefore, when moving to an FCoE environment, a phased approach is probably more realistic. A phased approach fits into your existing network with very little change, initially allowing you to keep your legacy equipment while realizing many of the benefits of FCoE. Consequently, the test topology we used was representative of a migration environment that most companies will initially move to when considering an FCoE environment. In this converged environment, the LAN and SAN shared the same infrastructure, enabling flexible access and improved availability for the storage.

For the tests, we decided to keep the non-network related elements of the design consistent to allow a direct comparison of the different network solutions, rather than focusing on the differences associated with server and storage hardware and operating systems. Therefore server hardware was common for all tests:

Server A

- HP DL380 G5
- Qlogic QLE8042 first generation Converged Network Adapter (CNA) and Host Bus Adapter (HBA) QLE2460

Server B

- HP DL380 G5
- Emulex LPE21002 first generation CNA and HBA LPE11000

At the time of testing, only first generation CNAs were available. Because of their architecture, these CNAs are currently restricted to a maximum FCoE throughput of 4Gbps per 10GE interface (as reflected in the test results). Future, second generation CNAs, already announced by both vendors will not have this limitation, and will be able to forward FCoE traffic at up to 10Gbps.

The networking elements used were as follows:

- VMware VSphere 4.0 as the host hypervisor (selected because of its integrated support for FC, FCoE, iSCSI and NFS). We also felt that this represented one of the more probable initial deployment scenarios for consolidated I/O based on 10GE and FCoE technologies. Current VMware network connectivity best practice recommends multiple independent I/O connections for the various production, VMKernel, Service Console and SAN links, and is therefore ideally placed to benefit from I/O consolidation.
- Software versions

Hardware / Products	Software Version
Nexus 5020	NX-OS 4.0(1a)N2(1)
Nexus 2148T	NX-OS 4.0(1a)N2(1)
MDS 9513	NX-OS 4.1.(3a)
NetApp FAS3170	ONTAP 7.3.1P3D1
Proliant DL380 G5	ESX 4.0 Vsphere

Devices used:

- In the core, Nexus 7000 switches.
- At the edge of the network (access layer), an FCoE-capable Nexus 5000 switch, which provided IP connectivity for the servers, as well as target and initiator connectivity for the SAN. By including the Nexus 5000 switches, we achieved the I/O consolidation capability necessary for a completely unified networking environment.
- Nexus 2000 switches to provide support for 1GE connections.
- FCoE-capable NetApp 3170 controller for all storage (as it is the only solution currently available that supports all network options concurrently). We attached the NetApp 3170 directly to the Nexus 5000 via FCoE, and to the MDS 9513 via FC. This greatly enhanced the topology of the network and the available storage paths over those of a traditional network.

Figure 1 is a diagram of the test topology used for the NetApp/FCoE series of tests described in this document. In the diagram, green lines represent FC, red lines are Ethernet, purple lines are FCoE, and yellow lines are 1GE.

Figure 1 - NetApp/FCoE Test Topology: Diagram

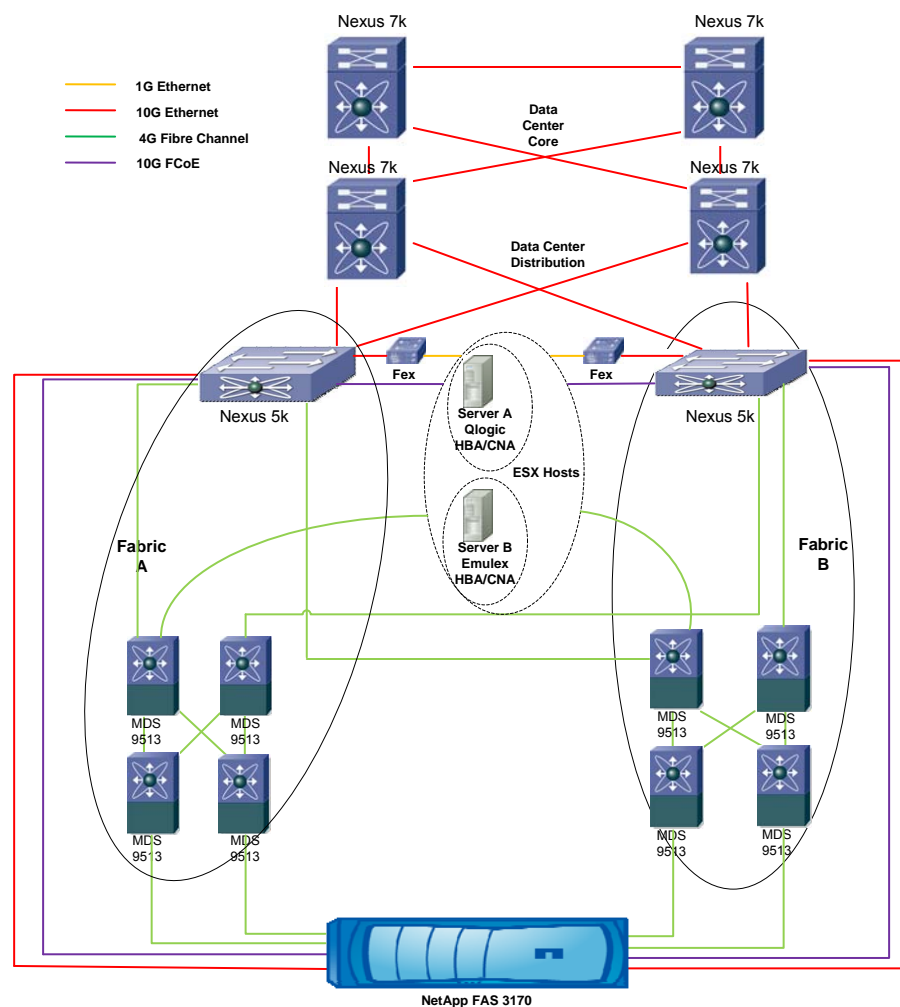
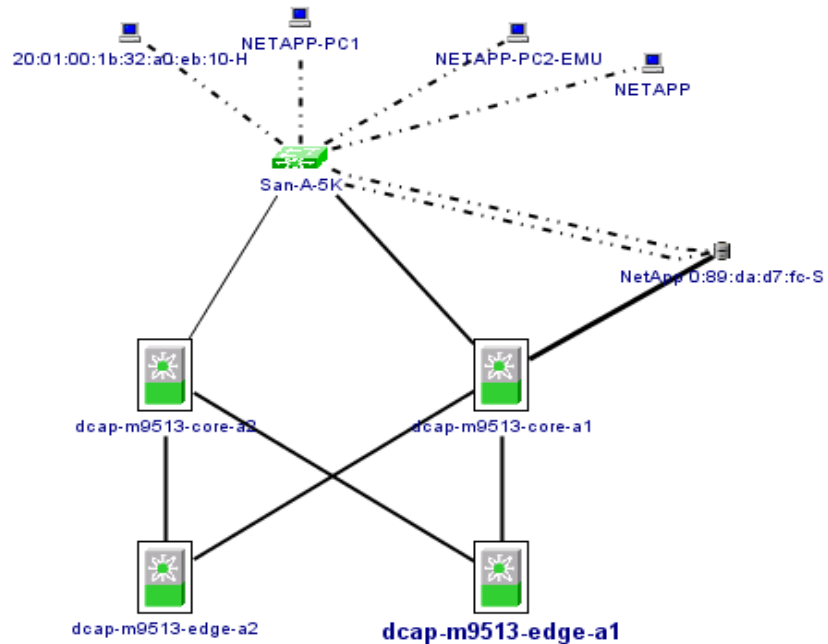


Figure 2 shows this unified fabric configuration as seen with the Cisco Fabric Manager interface. The Nexus 5000 switches are managed by Fabric Manager, and are visible as part of the existing SAN network, providing seamless continuity to the data center environment. The dotted lines represent the FCoE connections.

Figure 2 - NetApp/FCoE Test Topology: Fabric Manager View



FCoE Configuration Guidelines and Considerations

When deploying a fabric using FCoE, there are some requirements and configuration considerations to be noted.

Licensing

On the Cisco Nexus 5000, FCoE support is provided by the Storage Protocol Services License. The correct licenses are N5010SS or N5020SS and are activated using the **feature fcoe** command.

On the NetApp FAS, FCoE support is enabled through the FC license at no additional cost. Similarly, iSCSI support is provided through the NFS license.

Networking

FCoE requires the following networking features:

- Lossless Ethernet. The switches must support PAUSE frames, and provide “no-drop behavior” for FCoE traffic – i.e., must not discard FCoE frames due to lack of forwarding or buffer resources.
- Support for “Baby” Jumbo Frames (frames up to 2500 bytes in size) in order to avoid segmentation of FCoE frames containing the maximum sized FC frame of 2148 bytes .
- IEEE Data Center Bridging technologies (DCB) such as Priority Flow Control (PFC), Enhanced Transmission Service (ETS) and Data Center Bridging Capability Exchange

(DCBX) provide additional functionality designed to optimize the performance of converged I/O traffic in a Unified Fabric. These capabilities are present with the Nexus 5000 switches on 10GE interfaces.

Hardware Adapters

Converged network adapters (CNAs) are 10GE network interface cards (NICs) that function as both network adapters and storage adapters. For our testing, we used off-the-shelf CNAs from Qlogic and Emulex. Installation of both Qlogic and Emulex CNAs on the host server requires the use of the PCIe bus. When installed on ESX 4.0, the operating system comes with the drivers pre-loaded and the HBA/CNAs are recognized. The 10GE CNA interfaces appear in the ESX host as separate network adapters and storage adapters. These CNAs also support PAUSE/PFC flow control, “Baby” Jumbo Frames and DCB ETS and DCBX functions.

There is no major difference between the installation of an HBA and a CNA. However using a CNA automatically gives you an increase in LAN bandwidth and the potential to reduce the number of NIC adapters in your server.

Figure 3 – CNAs Storage Interfaces (FC Connectivity View) as shown in the ESX/VCenter

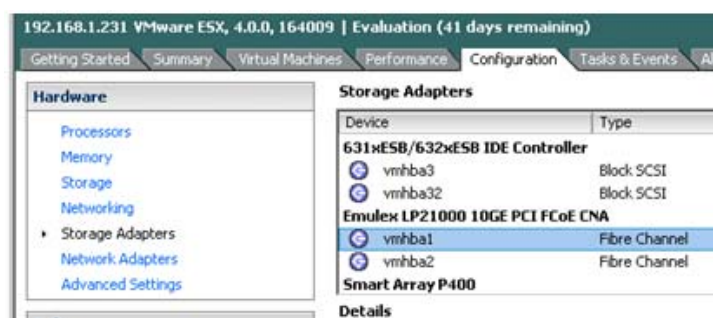


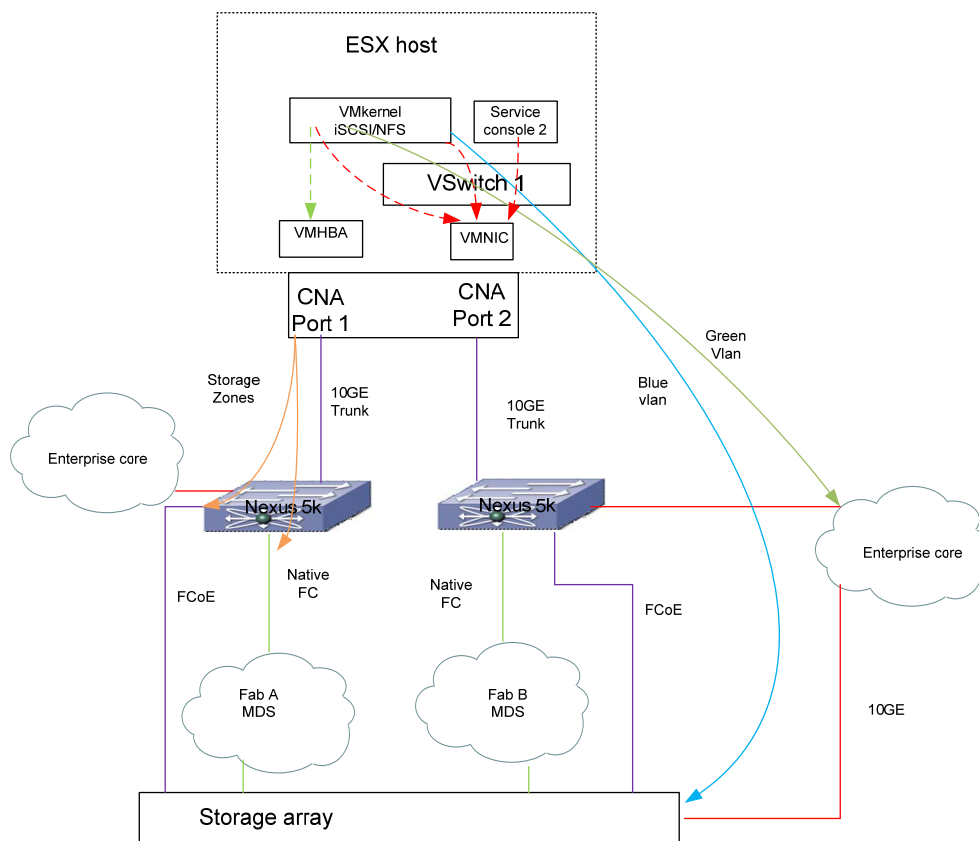
Figure 4 – CNA 10GE Interfaces (IP Connectivity View) as shown in ESX/VCenter

Device	Speed	Configured	Switch	MAC Address
Broadcom NetXtreme II BCM5708 1000Base-T				
vmnic1	100 Full	100 Full	vSwitch0	00:1c:c4:ee:d6:3a
vmnic0	1000 Full	1000 Full	None	00:1c:c4:ee:d6:3c
82598EB 10 Gigabit AF Dual Port Network Connection				
vmnic3	10000 F...	10000 Full	vSwitch1	00:00:c9:76:f3:6b
vmnic2	10000 F...	10000 Full	vSwitch1	00:00:c9:76:f3:6a

We also used Qlogic CNAs in the NetApp arrays, to allow the array to be connected directly to the Nexus 5000 switches. When configuring the network, consideration must be given to the hosts and VLANs which will be trunked. This can have implications for the cluster that may be created, the use of management networks, or the use of advanced functions such as VMotion.

Hosts and virtual machines have a number of connectivity choices when using FCoE to access storage in a unified fabric environment across the same adapter. One of these choices is iSCSI/NFS, as the servers still have IP connectivity. This can be implemented by allowing the network to which the array is attached to be carried directly to the host and to the VMkernel via the CNA 10GE trunk - alleviating the need for any routing.

Figure 5 – CNA Trunking

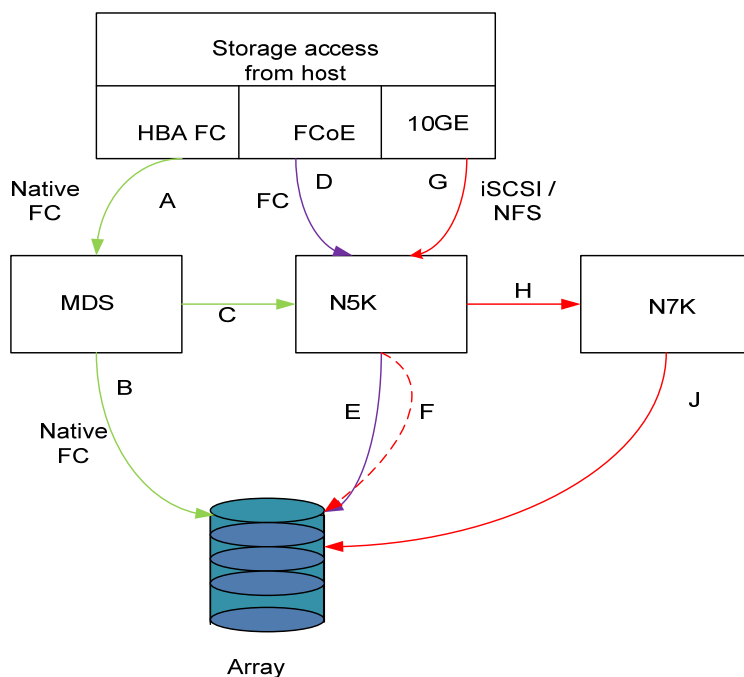


In addition, FC access has a number of paths via the CNA using the Zoning ability of Cisco Fabric Manager. Figure 5 (above) shows an FCoE-only Zone where the FCoE interface from the host can be zoned directly to the FCoE interface going to a directly attached native FCoE array (such as NetApp). There is also an FCoE connection from the host to native FC on the SAN (MDS 9000) network. In the diagram, green lines represent FC, red lines are 10GE, purple lines are FCoE, and yellow lines are 1GE. Dashed lines are virtual connections.

Available storage Zones in this unified environment are listed below, are shown in Figure 6. In the figure, green lines represent FC connections, red represents Ethernet, and purple represents FCoE. The following paths are shown in Figure 6.

- Block storage via HBA Path A-B,A-C-E
- Block storage via FCoE Path D-E, D-C-B
- NAS access via FCoE Adapter G-H-J, future(G-F)

Figure 6 – Available Storage Zones



Configuration of HBA on MDS:

```
interface fc1/1
  switchport speed 4000
  switchport rate-mode dedicated
  switchport mode F
  no switchport trunk allowed vsan all
  switchport trunk mode off
  no shutdown
```

```
dcap-m9513-edge-a2# sh flogi data
```

INTERFACE	VSAN	FCID	PORT NAME	NODE NAME
fc1/1	101	0x670004	10:00:00:00:c9:65:15:fc	20:00:00:00:c9:65:15:fc [NetApp-Hba-pc3]

Configuration of CNA on Nexus:

```

description PC2
no shutdown
bind interface Ethernet1/39

interface Ethernet1/39
description PC2 emulex CNA port 1
switchport access vlan 91
switchport trunk allowed vlan 91
fcoe mode on

vlan 91
fcoe vsan 101

vsan database
vsan 101 interface vfc11
vfc11 is up
Bound interface is Ethernet1/39
Port description is PC2
Hardware is GigabitEthernet
Port WWN is 20:0a:00:0d:ec:b1:9d:3f
Admin port mode is F
snmp link state traps are enabled
Port mode is F, FCID is 0x680006
Port vsan is 101
Beacon is turned unknown
5 minutes input rate 552 bits/sec, 69 bytes/sec, 0 frames/sec
5 minutes output rate 760 bits/sec, 95 bytes/sec, 0 frames/sec
12229247 frames input, 24518038000 bytes
0 discards, 0 errors
1186868 frames output, 359245720 bytes
0 discards, 0 errors

San-A-5K# sh flogi data
vfc11      101      0x680006   10:00:00:00:c9:76:f3:68   20:00:00:00:c9:76:f3:68
                                     [NetApp-PC2-EMU-cna]
```

Tests and Results

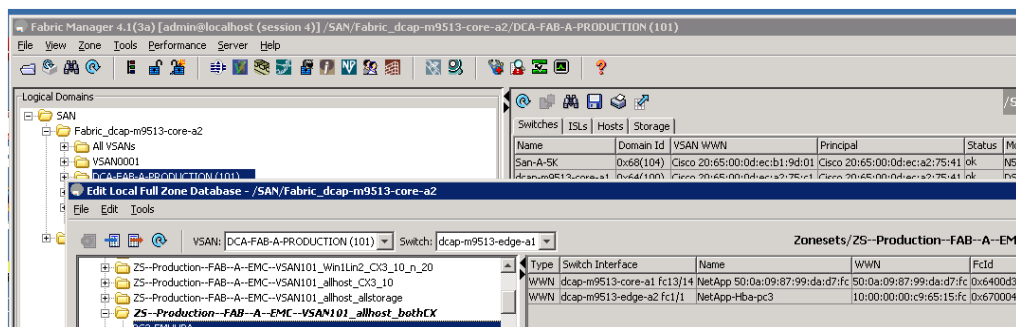
The testing was carried out using both Emulex and Qlogic CNAs and HBAs on ESX 4.0 hosts. Configuration to the NetApp 3170 array from the unified fabric was made using all available paths except path F, which only supports FCoE and no other Ethernet-based protocols. The comparison in configuration paths was made to highlight any practical significant differences in introducing FCoE into the network. The following tests were run:

- Zoning on the Unified Fabric
- Assigning Storage to Virtual Machines
- Performance of Adapters
- Priority Flow Control

Zoning on the Unified Fabric

Figure 7 shows the view of the fabric as seen using Fabric Manager. The Emulex HBA is in the ESX host PC3, which is zoned to the NetApp storage array via the ports on the MDS utilizing the FC-FC Route (A-B).

Figure 7 – Fabric Manager – Edit Local Full Zone Database



In Figure 8, the HBA ports are zoned to the NetApp array on the Nexus 5000 which is connected on FCoE interfaces via VFC4. This is to demonstrate the flexibility of a transitional topology which will allow the FC-FC Route (A-B) across a traditional SAN and FC-FCoE (A-C-E) connection. In such a connection, the initiators can remain connected to the SAN fabric with the HBA and still reach a native storage array on the Nexus 5000 using FCoE.

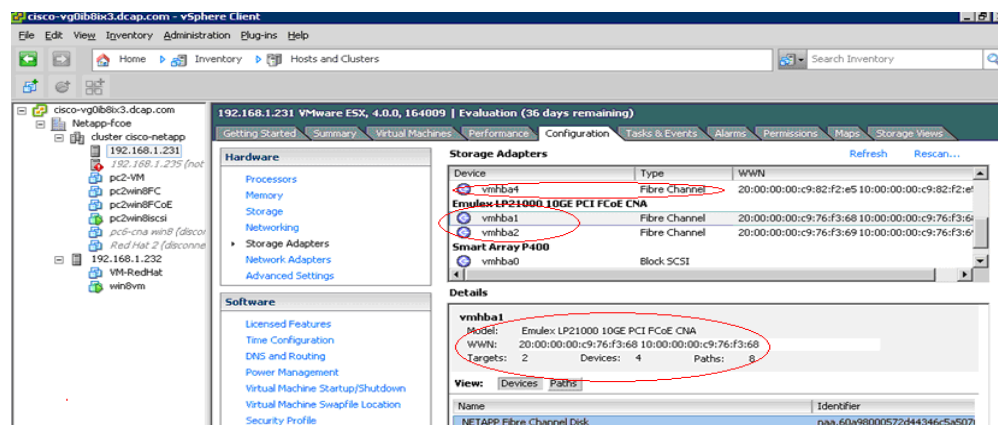
Figure 8 – Fabric Manager View of HBA Ports Zoned to NetApp Array



From a virtualization perspective the VMware hosts see both routes to the array, one via the HBA and the other via FCoE. This enables an easy migration from a legacy I/O environment based on GE & FC, to a consolidated I/O solution based on 10GE with FCoE. All data stores and RDM volumes can be accessed seamlessly via both FC and FCoE paths.

In Figure 9, the HBAs and CNAs can be seen, and the operating system has automatically identified the various paths to the available storage, in an environment where there had previously only been HBAs. This allows a migration path as we move towards a consolidated environment.

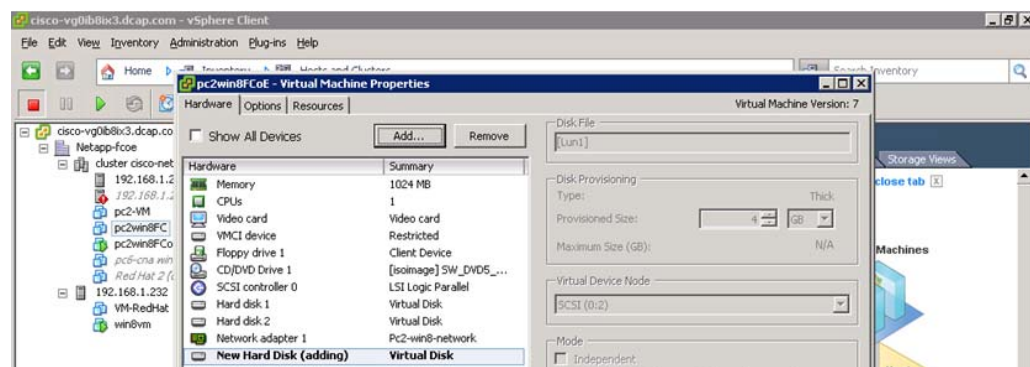
Figure 9 – Fabric Manager View of the HBAs and CNAs



Assigning Storage to Virtual Machines.

According to the testing, once the storage LUN is visible to the host, there is no impact to the virtual machines in assigning the storage, regardless of the path used. The storage is handled in exactly the same way via FCoE as it was previously via HBA. The operating systems tested on the virtual machines included Red Hat Enterprise Linux (RHEL) and Windows 2008.

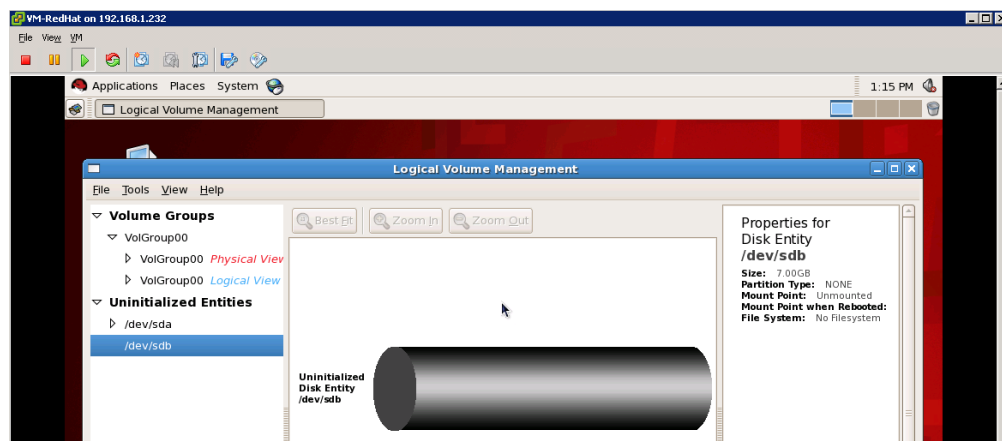
Figure 10 – Assigning Storage to Virtual Machines



The assignment of the storage in the operating system is independent of the adapters used.

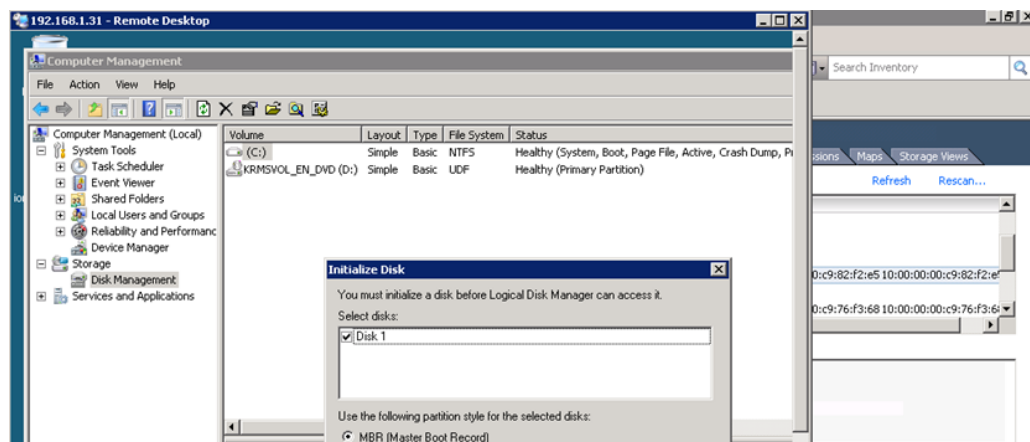
For A Red Hat virtual machine, the disk is initialized using the Logical Volume management as shown in Figure 11.

Figure 11 – Initialization of Storage using Logical Volume Management (Linux)



For Windows virtual machines, the new storage is initialized using the Disk Management menus as shown in Figure 12.

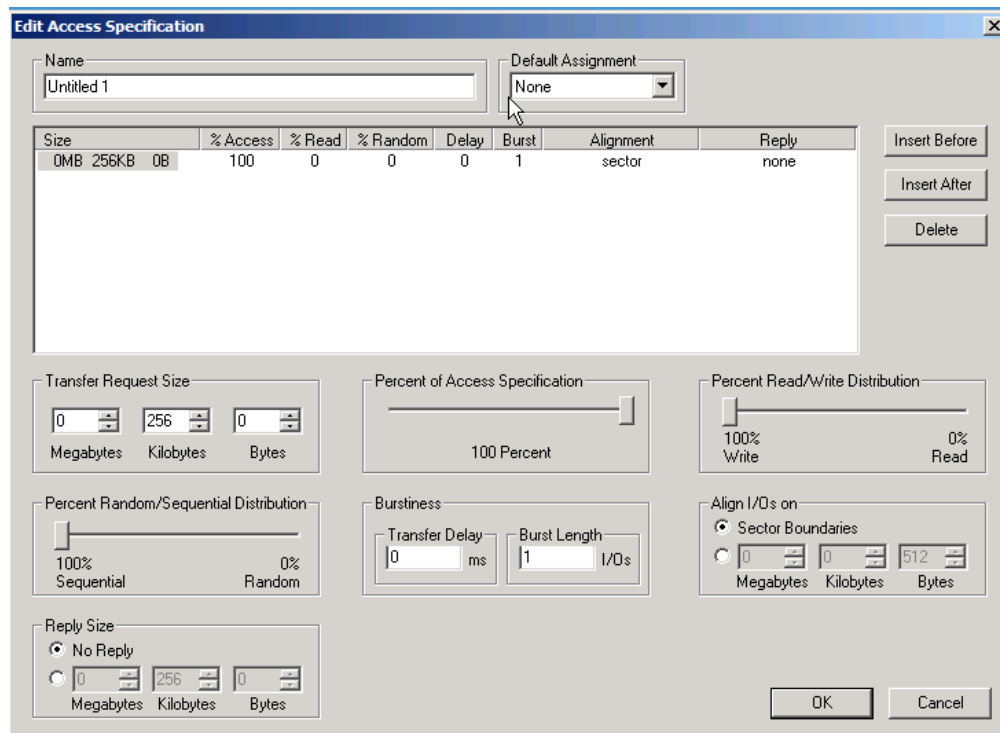
Figure 12 – Initialization of Storage using Disk Management (Windows)



Performance of Adapters

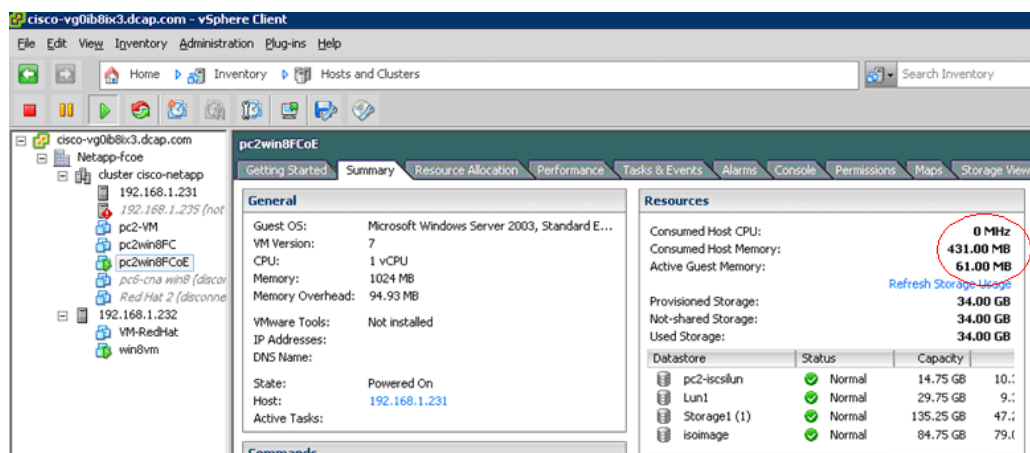
As part of the testing, a comparison of HBAs and CNAs was made to establish whether the adapters increased CPU utilization, or if there was any degradation in speed when using the CNA as opposed to the HBA. For this test, a Windows 2008 Server Virtual Machine was created and the IOMeter disk performance test tool was installed. A workload generator was created to generate read/write traffic capable of line rate operation at 4Gbps, as shown in Figure 13.

Figure 13 – IOMeter Workload Generator



Prior to running the IOMeter disk performance testing tool, a baseline for consumed resources was established, as shown in Figure 14.

Figure 14 - VMware Resource Utilization without IOMeter Tester Running



CNA (FCoE) Testing

For the next test, the NetApp LUNs were made accessible via the CNA (FCoE) paths and HBA (FC) paths. Also iSCSI was used via the CNA (10G Ethernet) adapter. We used the IOMeter test application to perform read/write operations on the disk and the port speed was noted. Using the CNAs, we were able to attain nearly the 4Gbps (400MBps) throughput, as recorded from the Windows 2008 virtual machine. Figure 15 shows the performance dial while the test was underway.

Figure 15 – IOMeter display (MB per second) during FCoE Tests



With Cisco Fabric Manager it was possible to verify the performance measurement using the Device Manager Summary View for Nexus 5000 switches attached.

When the IOMeter test was running, there was an observable increase in the virtual machine Consumed Host CPU resources. This was entirely due to the workload associated with the IOMeter application and the read/write processing to the disk.

HBA (FC) Testing

VMware was configured to access the test LUN via the Emulex HBA and the same IOMeter workload test was performed. Figure 16 shows the performance dial for this test.

Figure 16 - IOMeter display (MB per second) during FCoE Tests



The FC tests used less CPU but more memory than the FCoE tests. It is likely that this tradeoff between memory and CPU utilization is as a result of current driver implementations for the 1st Generation CNAs. It is anticipated that as FCoE drivers are optimized and 2nd generation hardware becomes available FC and FCoE resource utilization for FC and FCoE would be virtually identical at the same workloads.

iSCSI Tests

The same test was performed but using iSCSI via the CNA (10GE) interfaces. Performance levels similar to the FCoE and FC tests were obtained, but there was an increase in both CPU and memory overhead when using iSCSI.

As anticipated, the use of iSCSI which utilizes TCP/IP as its underlying transport consumes considerably more CPU and memory resources than either FC or FCoE.

Summary of Results

The following table summarizes the test results for the previous sections.

Summary of CPU and Memory Performance Tests on the ESX Host				
	IOMeter Performance (Workload) (MB/s)	Consumed CPU (MHz)	Active Guest Mem (MB)	Consumed Host Mem (MB)
VM Idle no test running	-	23	194	494
Emulex FCoE CNA (during test)	387.65	956	51	431
Idle State (after test)	0	956	61	431
Emulex FC HBA (During test)	387.65	419	71	568
NFS (10GE) CNA (During test)	357.27	2029	174	849
iSCSI (10GE) CNA (During test)	387.65	1166	163	908
iSCSI (1GE) LOM* (During test)	109.96	583	102	657
Qlogic FCoE CNA (During test)	384.28	1189	51	771
Qlogic FC HBA (During test)	383.69	1166	51	935

*LOM=LAN on motherboard

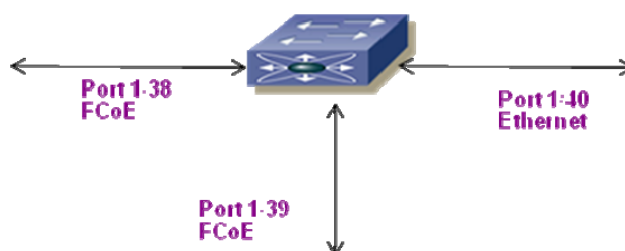
Note: The data above is for implementation comparison purposes only based on the single host test environment used.

Priority Flow Control

One of the concerns storage administrators have about using FCoE is if the FC traffic will be affected and storage suffer as a result, when IP/TCP traffic is attempting to use more than available bandwidth. Testing has shown that with a very basic configuration on the Nexus 5000, this is not a concern. FC traffic will not be dropped if there is an increase in IP traffic. The Nexus 5000 implements priority flow that protects FC traffic. This test was done in isolation of the network using an Ixia tester connected to one Nexus 5000 as the DUT. All traffic was generated and terminated on the Ixia without the use of an array.

In the topology and example below, three IXIA ports are connected to a Nexus 5020 switch. Two ports are FCoE ports and one port is IP. Both FCoE ports have successfully logged onto the fabric.

Figure 17 – Topology: Three IXIA Ports connected to a Nexus 5020 Switch



Example

```
dcap-SANB-edge-C3# sh flogi database
```

INTERFACE	VSAN	FCID	PORT NAME	NODE NAME
vfc5	301	0x120001	21:00:00:1b:32:80:ef:0f [sanB-Netapp-PC5]	20:00:00:1b:32:80:ef:0f
vfc99	301	0x120003	fb:13:00:11:0d:01:00:00	fb:13:00:11:0d:01:00:00
vfc100	301	0x120004	fb:12:00:11:0d:01:00:00	fb:12:00:11:0d:01:00:00

Total number of flogi = 3.

```
dcap-SANB-edge-C3#
```

Configuration

```
feature fcoe
vlan 81
  fcoe vsan 301
vsan database
  vsan 301 name "DCB-FAB-C-PRODUCTION"
fcdomain fcid database
  vsan 301 wwn fb:13:00:11:0d:01:00:00 fcid 0x120003 dynamic
  vsan 301 wwn fb:12:00:11:0d:01:00:00 fcid 0x120004 dynamic
interface vfc99
  bind interface Ethernet1/38
  no shutdown
interface vfc100
  bind interface Ethernet1/39
  no shutdown
interface Ethernet1/38
  switchport mode trunk
  switchport trunk allowed vlan 81
  fcoe mode on
  priority-flow-control mode on
interface Ethernet1/39
  switchport mode trunk
```

```

switchport trunk allowed vlan 81
fcoe mode on
priority-flow-control mode on
interface Ethernet1/40
switchport access vlan 81
zoneset distribute full vsan 301
zone name Z-Net-Ixia vsan 301
  member pwn fb:12:00:11:0d:01:00:00
  member pwn fb:13:00:11:0d:01:00:00
zoneset name Z-Ixia vsan 301
  member Z-Net-Ixia
zoneset activate name Z-Ixia vsan 301

```

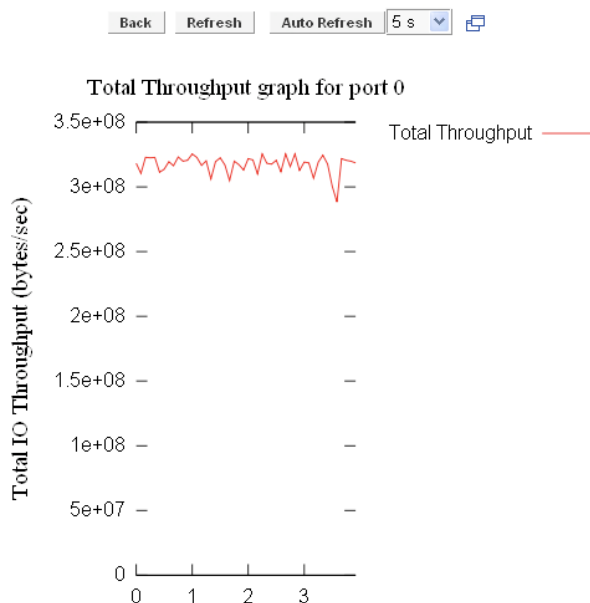
For the testing procedure, we sent the line rate legacy Ethernet traffic (10G) from the IXIA on port 1/40 to port 1/39, and recorded throughput for Ethernet only. We expected to get almost 10G. As shown in Figure 18 below, the sent rate is equal to the received rate of traffic.

Figure 18 – Send Rate equal to Received Rate

Bits Sent Rate	0	9,624,059,891
Bits Received	21,318,273,32	2,668,512
Bits Received Rate	9,626,610,730	982
Control Bits Transmitted (%)	44	42

When we started FC traffic from an FCoE connected port 1/38 (initiator) to port 1/39 (target), we got the baseline throughput with IxSAN.

Figure 19 – Throughput Graph for Port 0



The stats above show that the Ixia throughput is approx 320 Megabytes/second, which roughly equals 2.56 Gigabits/second. This is verified by the port statistics on the switch as shown in the configuration example below, by adding together the TX and RX figures to get the throughput value.

```

dcap-SANB-edge-C3# sh int vfc 99
vfc99 is up
  Bound interface is Ethernet1/38
  Hardware is GigabitEthernet
  Port WWN is 20:62:00:0d:ec:b2:bc:bf
  Admin port mode is F
  snmp link state traps are enabled
  Port mode is F, FCID is 0x120003
  Port vsan is 301
  Beacon is turned unknown
  5 minutes input rate 1675919848 bits/sec, 209489981 bytes/sec, 812856
frames/sec
  5 minutes output rate 890085968 bits/sec, 111260746 bytes/sec, 437704
frames/sec
    591751748 frames input, 1227610079332 bytes
    0 discards, 0 errors
    143774596 frames output, 282317558480 bytes
    0 discards, 0 errors

dcap-SANB-edge-C3# sh int vfc 100
vfc100 is up
  Bound interface is Ethernet1/39
  Hardware is GigabitEthernet
  Port WWN is 20:63:00:0d:ec:b2:bc:bf
  Admin port mode is F
  snmp link state traps are enabled
  Port mode is F, FCID is 0x120004
  Port vsan is 301
  Beacon is turned unknown
  5 minutes input rate 876950528 bits/sec, 109618816 bytes/sec, 435688
frames/sec
  5 minutes output rate 1694364568 bits/sec, 211795571 bytes/sec, 813912
frames/sec
    144309448 frames input, 280517452960 bytes
    0 discards, 0 errors
    592770393 frames output, 1241565292624 bytes
    0 discards, 0 errors

dcap-SANB-edge-C3#

```

When we sent both Ethernet and FC traffic together, Ethernet packets were dropped, as can be seen from the Ethernet statistics in Figure 20 below.

Figure 20 – Ethernet Statistics

Bits Sent	11,788,268,32	21,856,406,78
Bits Sent Rate	2,657,891,885	9,624,060,196
Bits Received	21,810,189,11	2,687,632
Bits Received Rate	7,040,468,844	0
Central Chip Temperature(°C)	44	43

FC traffic was error free according to the IxSAN detailed statistics below.

Figure 21 – IxSAN Detailed Statistics

Fibre Channel Counters	Total Since Reset	Avg Since Reset	Avg Since Last Read Less Detail
TxFrames	110510K	26469/sec	17806/sec
RxFrames	529842K	126908/sec	39467/sec
TxWords	5.19063e+07M	12432.6K/sec	8747.47K/sec
RxWords	2.69697e+08M	64598K/sec	19966.6K/sec
LIPCount	0	0/sec	0/sec
NOSCount	0	0/sec	0/sec
ErrorFrames	0	0/sec	0/sec
DumpedFrames	0	0/sec	0/sec
LinkFailure	0	0/sec	0/sec
LossOfSync	0	0/sec	0/sec
LossOfSignal	0	0/sec	0/sec
PrimitiveSequenceErrors	0	0/sec	0/sec
InvalidTxWord	0	0/sec	0/sec
InvalidCRCCount	0	0/sec	0/sec
FCPIInitiatorIOCCount	0	0/sec	0/sec
Time Since Reset	4175 seconds (0 days, 1 hour, 9 minutes, 35 seconds)		ClearCounters ReadCounters
Update Interval	<input type="text" value="0"/> <input type="button" value="v"/> Seconds <input type="button" value="Start Auto Refresh"/>		

The statistics below show that the TX and RX figures on both vFC interfaces are the same, and there are no errors. This proves that the FC traffic was still flowing without dropping packets.

```

dcap-SANB-edge-C3# sh int vfc 99s
vfc99 is up
  Bound interface is Ethernet1/38
  Hardware is GigabitEthernet
  Port WWN is 20:62:00:0d:ec:b2:bc:bf
  Admin port mode is F
  snmp link state traps are enabled
  Port mode is F, FCID is 0x120003
  Port vsan is 301
  Beacon is turned unknown
  5 minutes input rate 1675919848 bits/sec, 209489981 bytes/sec, 812856
frames/sec
  5 minutes output rate 890085968 bits/sec, 111260746 bytes/sec, 437704
frames/sec
    591751748 frames input, 1227610079332 bytes
      0 discards, 0 errors
    143774596 frames output, 282317558480 bytes
      0 discards, 0 errors

dcap-SANB-edge-C3# sh int vfc 100
vfc100 is up
  Bound interface is Ethernet1/39
  Hardware is GigabitEthernet
  Port WWN is 20:63:00:0d:ec:b2:bc:bf
  Admin port mode is F
  snmp link state traps are enabled
  Port mode is F, FCID is 0x120004
  Port vsan is 301
  Beacon is turned unknown
  5 minutes input rate 876950528 bits/sec, 109618816 bytes/sec, 435688
frames/sec
  5 minutes output rate 1694364568 bits/sec, 211795571 bytes/sec, 813912
frames/sec
    144309448 frames input, 280517452960 bytes
      0 discards, 0 errors
    592770393 frames output, 1241565292624 bytes
      0 discards, 0 errors

```

Conclusion and Recommendations

The purpose of these tests was to establish whether FCoE could provide a practical alternative to FC, by assessing and comparing the operational and performance characteristics of the two technologies in a typical virtual server environment. The rationale for using FCoE to facilitate I/O consolidation is discussed at length in numerous other white papers and case studies, but to summarize, consolidated I/O can:

- Dramatically reduce the number of I/O connections per server, especially in Virtual Machine environments
- Significantly reduce the number of switch ports per data center
- Eliminate the need to run multiple network technologies inside the DC
- Greatly simplify the access layer cabling for data centers

All these can deliver genuine and significant capital and operational expense reductions for data centers.

The testing also sought to compare FCoE to other commonly used options for achieving I/O consolidation (i.e. protocols allowing storage access via Ethernet) in use today, by looking at how both iSCSI and NAS solutions perform and are provisioned versus FCoE and FC. In both cases, NAS and iSCSI were seen to consume greater host resources than either FC or FCoE and required different operational practices for provisioning and monitoring. Quite clearly, iSCSI and NAS are powerful technologies and offer flexible solutions, but currently they seem more suited to environments where workloads are at the lower end of the spectrum. In terms of performance versus host resource utilization, FC and FCoE seem to have an advantage.

However, FCoE will only be acceptable if it performs at least as well (if not better) than FC, yet can be provisioned, managed and monitored using the same basic tools and procedures as FC. The testing and associated results detailed in this document demonstrate the FCoE has performance and resource utilization characteristics similar to FC, even using 1st generation FCoE CNA technology. The advent of 2nd generation FCoE CNAs will allow for throughput for FCoE traffic up to the full line rate of the 10GE interfaces (approximately 1200MB/s). This has been verified by internal testing for pre-production 2nd generation CNAs. Likewise, the tests demonstrate that the volume provisioning, fabric zoning and host LUN provisioning are identical across both FC and FCoE environments, and the same tools can be used to manage both solutions. We were also able to show that performance monitoring of SAN traffic was the same across both FC and FCoE networks. Based on these results, storage and server managers should have no concerns about the performance and operational aspects of deploying FCoE versus FC.

Additional testing also demonstrated that the fundamental elements underlying the use of FCoE, such as Per-Priority Flow Control, work reliably and consistently, delivering the lossless fabric interconnect needed by SAN traffic.

The topology options associated with FCoE deployments are somewhat restrictive (i.e., the lack of switch VE_Port implementations means that only a single switch hop is currently supported for FCoE traffic). However, it has been shown that in these environments FC and FCoE perform almost identically. Future enhancements to FCoE products such as FCIP Virtual E Port (VE_Port) support and FCoE Initialization Protocol (FIP) snooping capabilities will enable much more complex topologies to be built, at least comparable (and ultimately larger and more flexible) than current FC fabrics. Even with the current topology limitations, FCoE provides a very effective means to rationalize the server interfaces for LAN and SAN connections into a single pair of 10GE interfaces using 1st generation CNAs, while 2nd generation CNAs will offer greater performance and flexibility

with both reduced power consumption and smaller form factor, making them suitable for a wide variety of server platforms. Interoperability with existing FC SANs means that all current storage is accessible via the 10GE consolidated I/O interfaces. However, it is already possible to completely eliminate the need for FC fabrics in simple environments, as NetApp controllers now support native FCoE interfaces. While current limitations require that these be connected to the same switches as the servers (i.e. single switch hop between server and storage), it is now quite feasible to build solutions entirely based on FCoE over 10GE from end-to-end – without the need for any FC components. As these topology restrictions are removed with the advent of VE_port and FIP snooping, the capital and operational expense saving associate with unified fabrics will mean that FCoE (along with 10GE) will become the dominant technology for SAN connectivity.

In order to reap the considerable benefits associated with consolidated I/O and unified fabrics, we would recommend that anyone currently planning or building next generation data centers, factor this technology in to the design process. We would also suggest that anyone currently planning additional capacity or upgrading existing data centers carefully look at FCoE technology today as a possible option for server I/O consolidation, or as an end-to-end solution for simple single stage fabrics.



Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Europe Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: +31 0 800 020 0791
Fax: +31 0 20 357 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

©2006 Cisco Systems, Inc. All rights reserved. CCVP, the Cisco logo, and the Cisco Square Bridge logo are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Registrar, Aironet, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, GigaStack, HomeLink, Internet Quotient, IOS, IP/TV, IQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, LightStream, Linksys, MeetingPlace, MGX, Networking Academy, Network Registrar, Packet, PIX, ProConnect, RateMUX, ScriptShare, SlideCast, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0609R)