Customer Benefits of Using Converged Network Adapter (CNA) – A Case Study

March 2014

Saritha Vinod and Anto A John

IBM Systems and Technology Group

1. Overview

With the rapid development of today's data centers and the increased number of servers deployed, independent deployment of LAN (Local Area Network) and SAN (Storage Area Network) results in various problems like increased cost and complex infrastructure management. Hence protocol standardization bodies and network and storage product vendors have been working on convergence of storage and network traffic in a data center. The term “convergence” here refers to consolidation of storage traffic and traditional data traffic on a single network (or fabric).

Several factors are driving the interest in convergence of storage and network like faster transmission speeds of Ethernet, exceeding or meeting the speed available for Fibre Channel (FC) and enhanced capabilities of Ethernet making it low latency and lossless to accommodate different traffic types like FC.

Another major motivation for storage and networking convergence is improved asset utilization and cost of ownership. By using a single infrastructure for multiple types of network traffic, the costs of procuring, installing, managing, and operating the data center infrastructure can be lowered. In data centers where multiple types of adapters, switches, and cables were once required for separate networks, a single set of infrastructure will take its place, providing savings in equipment, cabling, space and power requirements. The improved speeds and capabilities of a lossless 10-Gbps Ethernet offer the hope of such improvements.

This paper covers the various technologies and protocols that are used to achieve this convergence and their advantages. We study the benefits of using converged network adapter using different scenarios and compare it against traditional setup. We elaborate on our test results and discuss various inferences out of it.

1.1 Converged Network Adapter (CNA)

A Converged Network Adapter (CNA) converges the functionality of a Storage Host Bus Adapter (HBA) and a Network Interface Controller (NIC) into one adapter. The same adapter can connect servers to storage area networks (SANs) and Ethernet-based local area networks (LANs).

1.2 Fibre Channel Over Ethernet (FCoE) and Other Standards

The IEEE 802.1 Data Center Bridging (DCB) Task Group created a set of standards that enhance existing 802.1 bridge definitions. The enhancements provide a converged network that allows multiple applications to run over a single physical infrastructure. A standardization activity for FCoE (Fibre Channel Over Ethernet) started in April 2007 inside the Fibre Channel - Backbone - 5 (FC-BB-5) working group of T11. FCoE is a standard that encapsulates FC frames over the Ethernet packets to be transmitted on Ethernet Networks. It relies on a new Ethernet transport with extensions that provide lossless transmission of storage data. The goal is to provide I/O consolidation over Ethernet, reducing network complexity in the Data center.

The FC-BB-5 standard specifies that FCoE is intended to operate over an Ethernet network that does not discard frames in the presence of congestion. Such an Ethernet network is called a lossless Ethernet in this standard. In order to achieve this, FCoE makes use of the DCB standards like Priority-based Flow Control (PFC) and Enhanced Transmission Selection (ETS).

Some of the DCB standards like Priority-based Flow Control (PFC), Enhanced Transmission Selection (ETS), Congestion Notification (CN), and the Data Center Bridging Capabilities Exchange protocol are discussed below.

- **Priority-based Flow Control (PFC)** provides a link level flow control mechanism that can be controlled independently for each priority. The goal of this mechanism is to ensure zero loss due to congestion in DCB networks. PFC is not necessarily restricted to FCoE networks. It can be used for loss-sensitive traffic in any network where traffic is separated into different priorities. In an FCoE context, the FCoE traffic is mapped to the priority level on which the selective pause is enabled. Therefore, if the receiving switch becomes congested, it does not continue receiving traffic that must not be dropped.

- **Enhanced Transmission Selection (ETS)** provides a common management framework for assignment of bandwidth to traffic classes. With ETS, port bandwidth is allocated based on 802.1p priority values in the VLAN tag. Multiple priority values are combined into priority groups and different amounts of link bandwidth are assigned to these groups. If a traffic group does not use its allocated bandwidth, other traffic is allowed to use that bandwidth. ETS allows multiple types of traffic to coexist on a converged link without imposing contrary handling requirements on each other. Similar to PFC, ETS is not restricted to use with FCoE traffic. ETS can be used in general to allocate link bandwidth among different traffic classes. In an FCoE context, ETS is used to give the FCoE traffic a certain priority level and bandwidth allocation.

- **Congestion Notification (CN)** provides end to end congestion management for protocols that do not already have congestion control mechanisms built in; e.g. FCoE. It is also expected to benefit protocols such as TCP that do have native congestion management as it reacts to congestion in a more timely manner.

- **Data Center Bridging Capabilities Exchange (DCBX)** protocol is a capabilities-exchange protocol that is used by DCB-capable switches to identify and communicate with each other. Neighboring devices use DCBX to exchange and negotiate configuration information and detect misconfigurations. This protocol is expected to leverage functionality provided by IEEE 802.1AB (LLDP). Among other tasks, devices use DCBX to exchange information about FCoE, information about whether PFC is enabled on the port, and ETS priority group information, including 802.1p priority values and bandwidth allocation percentages.

2. Benefits of using CNA

Following are some of the major benefits of using converged network adapters:

- High-performance storage access over loss-less 10 Gigabit Ethernet fabrics
- Transparent access to storage devices using existing SAN (storage area network) management methods
- Retention of enterprise-proven Fibre Channel drivers and management tools
- Lower capital, energy and cooling costs with fewer adapters, cables and switches
- Lower management overhead by maintaining a single fabric
- Increased application availability by simplifying the network and server configuration

3. Case Study Objective - Convergence using FCoE

In a traditional data center, the LAN and SAN are deployed and maintained separately. Compared to this, using a FCoE CNA helps to simplify the management of LAN and SAN and it reduces costs and energy consumption due to fewer adapter cards, cables, switch ports, and PCIe slots. This section describes the use cases in detail.

In a traditional setup having both LAN and SAN separately, the server would have a separate NIC connected to an Ethernet switch and an FC HBA connected to a Fibre Channel switch. All the LAN traffic is routed through the NIC and Ethernet switch, while FC HBA and Switch handles the FC traffic. The below diagram shows the traditional setup with NIC and FC HBA.

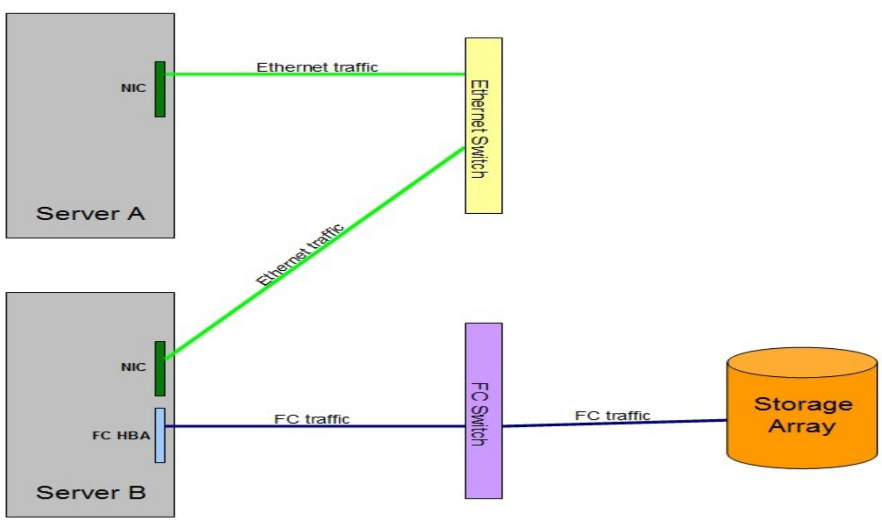

Fig 1: Traditional setup with FC HBA + Ethernet NIC on a server

The FCoE CNA connects to the server via a PCI Express (PCIe) interface. The server sends both FC SAN and LAN traffic to an Ethernet port on a converged switch using the FCoE protocol for the FC SAN data and the Ethernet protocol for LAN data. The converged switch converts the FCoE traffic to FC and sends it to the FC SAN. (Note: if a storage array with native FCoE support is used, the conversion from FCoE to FC at the switch level is eliminated.) The Ethernet traffic is sent directly to the LAN. This scenario is depicted in the figure below:

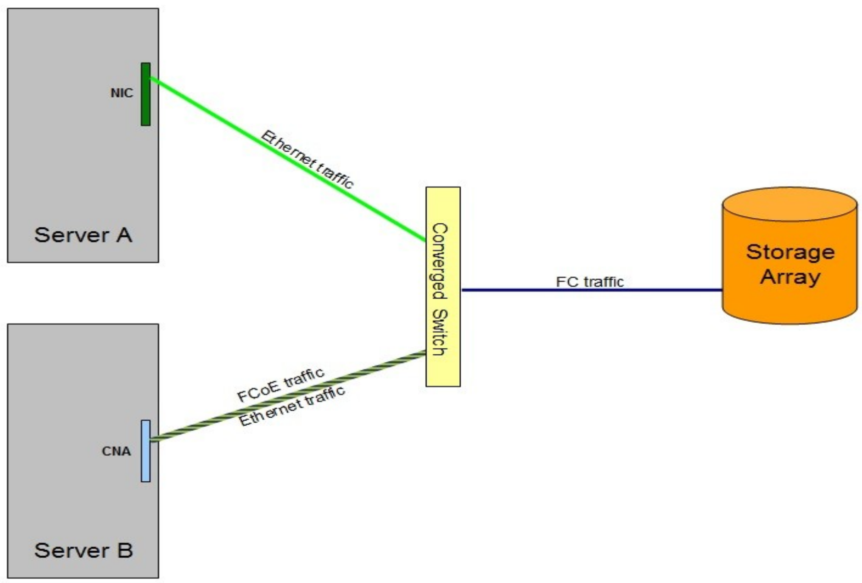

Fig 2: FCoE CNA on the server used for both FC and Ethernet Traffic

In section 5.1 of this white paper, we will be comparing the above two cases, thereby bringing out the benefits in using FCoE CNA for storage and network convergence. We also showcase the ability of the CNA to scale up the throughput with increase in the number of ports. Further, we portray the ETS capability for bandwidth allocation and resultant throughput variation of both Ethernet and FCoE traffic.

4. Test Setup and Configuration Overview

A Power7+ ITE (model 7895-43X) within a Flex System Chassis was used as the primary system on which the measurements were taken. The system had 16 Power7+ cores with SMT (Simultaneous Multi Threading) enabled. The server was running Linux operating system and the distro used was Red Hat Enterprise Linux Server release 6.4 (Santiago) with kernel version 2.6.32-358.el6.ppc64.

The measurements were done with a storage workload and network workload that was run simultaneously on the Power7+ server systems. The storage workload does a raw IO with sequential Read, Write and Duplex patterns and measures the throughput achieved in MB/s. The network workload does network IO with Receive, Transmit and Duplex patterns and measures the network throughput in MB/s. For the case study described in this whitepaper, both the workloads were run simultaneously with Read/Receive, Write/Transmit, Duplex/Duplex, Read/Transmit and Write/Receive patterns and the total throughput for the storage and network traffic is captured. Section 5 on Evaluation Results discusses the results and inferences.

Below sections detail the test setup and configuration of the two cases that was used for the study and comparison.

Case 1: Convergence of LAN and SAN using FCoE CNA

To showcase the convergence of LAN and SAN we created a setup with 10Gb FCoE converged network adapters (IBM Flex System CN4058 8-Port 10Gb Converged Adapter) on Power7+ systems used for performance measurements. A Power7+ client machine was also used as the client system while doing the network measurements. Both the server and client machines have two 10Gb FCoE adapters. Two ports from each of the 8-port FCoE adapters are used for 4-port testing measurements. The FCoE adapters are connected to two 10Gb Converged Switches (IBM Flex System Fabric CN4093 10Gb Converged Scalable Switch). Each port of the switch is able to transmit/receive either FC or Ethernet traffic. The converged switch routes the Ethernet packet to the specified destination in the LAN and the FCoE packets are converted to FC packets after/before transferring from/to the SAN. The setup is depicted below and the respective traffics are described through the legends in the diagram.

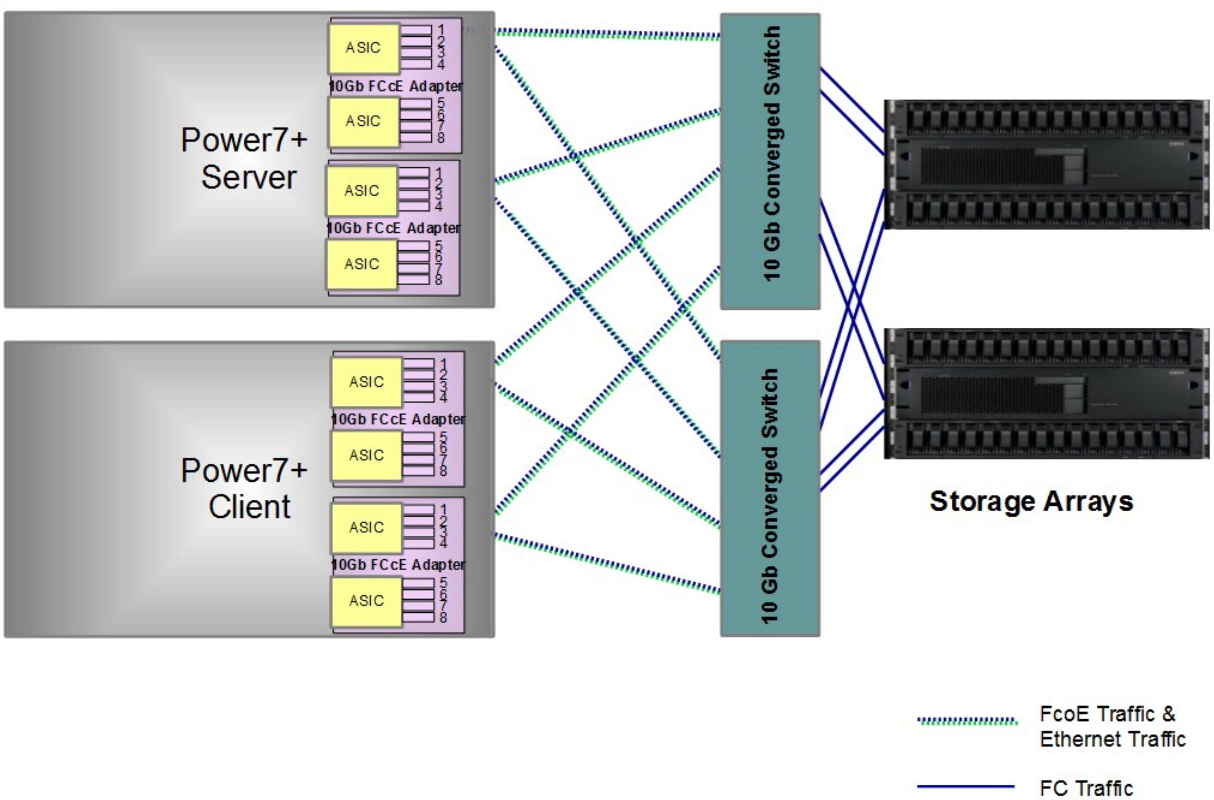

Fig 3: Test setup with 10Gb FCoE Converged Network Adapter

The back-end storage arrays used in this setup are IBM System Storage DS5300. Each DS5300 has two controllers (A and B). Only two host ports are connected to one DS5300 storage array to avoid any potential bottlenecks in accessing the back-end storage systems. Both the switches are zoned such that only one host port is associated with the utilized zones on a switch, so that the traffic is balanced across the switches. DS5300 is configured in such a way that switch 1 connects the ports to Controller A on DS5300 and switch 2 connects the ports to Controller B.

Case 2: Traditional setup with separate LAN and SAN

In contrast to the above, we also created another test setup to showcase a traditional use case where the server has both FC HBAs and Ethernet NICs to provide SAN and LAN connectivity respectively. In this case, the Power7+ server has two IBM Flex System FC3172 2-Port 8Gb FC Adapter, and two IBM Flex System EN4054 4-port 10Gb Ethernet Adapter, connected to it. The NICs are connected to an IBM Flex System Fabric EN4093 10Gb Scalable Switch and the FC HBAs are connected to an IBM Flex System FC5022 24-port 16Gb SAN Scalable Switch providing access to the SAN with DS5300 back-end storage systems. The DS5300 configuration is kept the same as in the above case.

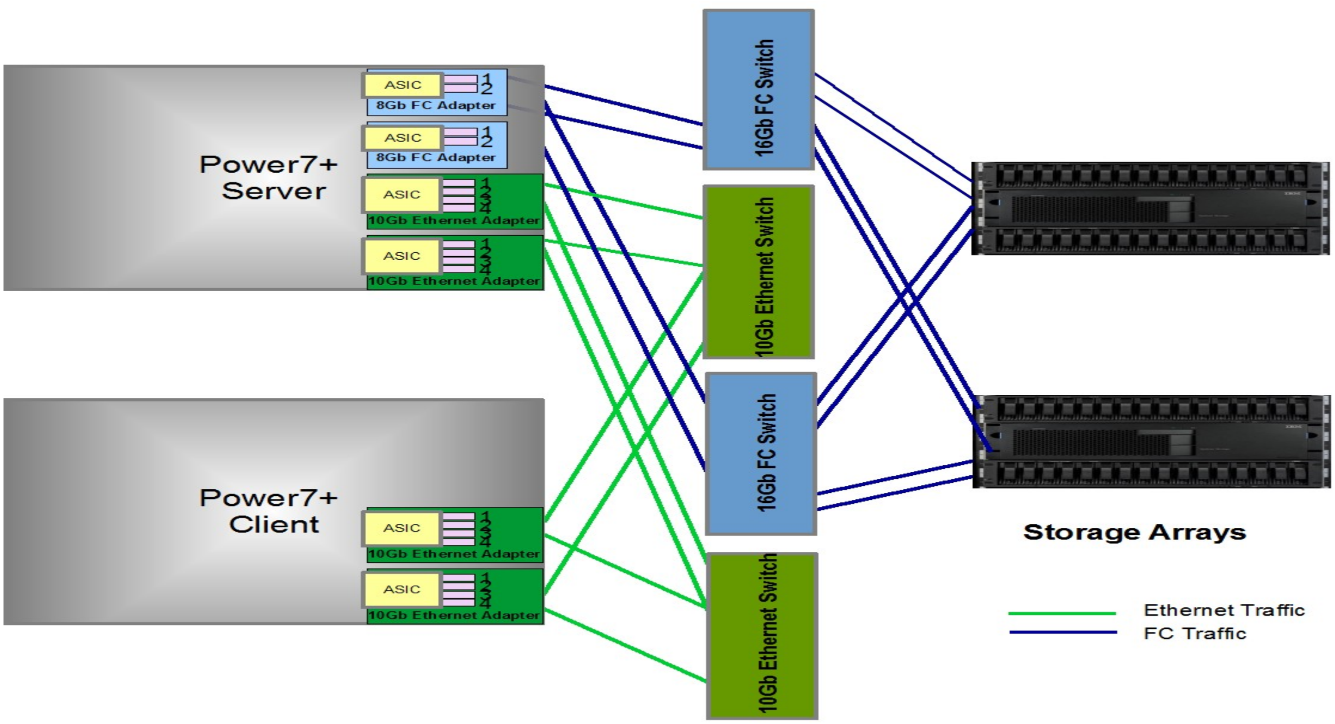

Fig 4: Test setup with 10Gb Ethernet Adapter and 8Gb FC Adapter

5. Evaluation Results

5.1 Bandwidth Utilization Comparison with FCoE CNA and Traditional FC HBA+NIC

In this section, we compare the bandwidth utilization of a traditional setup with separate FC and NIC adapters against the FCoE CNA.

In our test setup where we use separate FC and NIC adapters, the total theoretical bandwidth per port for LAN and SAN is 18Gbps from 8Gb FC and 10Gb Ethernet, whereas the FCoE CNA provides 10Gbps of theoretical maximum bandwidth per port. Since the total bandwidth capability of the two scenarios differ, normalization of the measured throughput was done using the below formulas (1) and (2) respectively, so that both the scenarios can be compared on the same scale.

Normalization formula for LAN + SAN scenario :

$$\lambda = \frac{x` + y`}{(x + y) \times N \times \beta} \qquad \text{------------------------------------- (1)}$$

Normalization formula for FCoE scenario :

$$\lambda = \frac{x` + y`}{Y \times N \times \beta} \qquad \text{------------------------------------- (2)}$$

In the above formulas (1) and (2),
λ stands for the normalized value
x stands for the theoretical maximum throughput of the NIC
y stands for the theoretical maximum throughput of FC HBA
Y stands for the theoretical maximum throughput of FCoE CNA
x' stands for measured throughput from Ethernet traffic
y' stands for measured throughput from FC traffic
N stands for the number of ports used
β stands for the duplex factor (full-duplex=1, half-duplex=0.5)

In the chart below, λ, the normalized value, is used for comparison of both the scenarios.

The chart clearly shows that the bus efficiency in case of FCoE CNA and FC + NIC case, are about equal. (NOTE: bus efficiency = (throughput/bandwidth), where throughput = actual work per unit of time and bandwidth = maximum theoretical work per unit of time).

λ

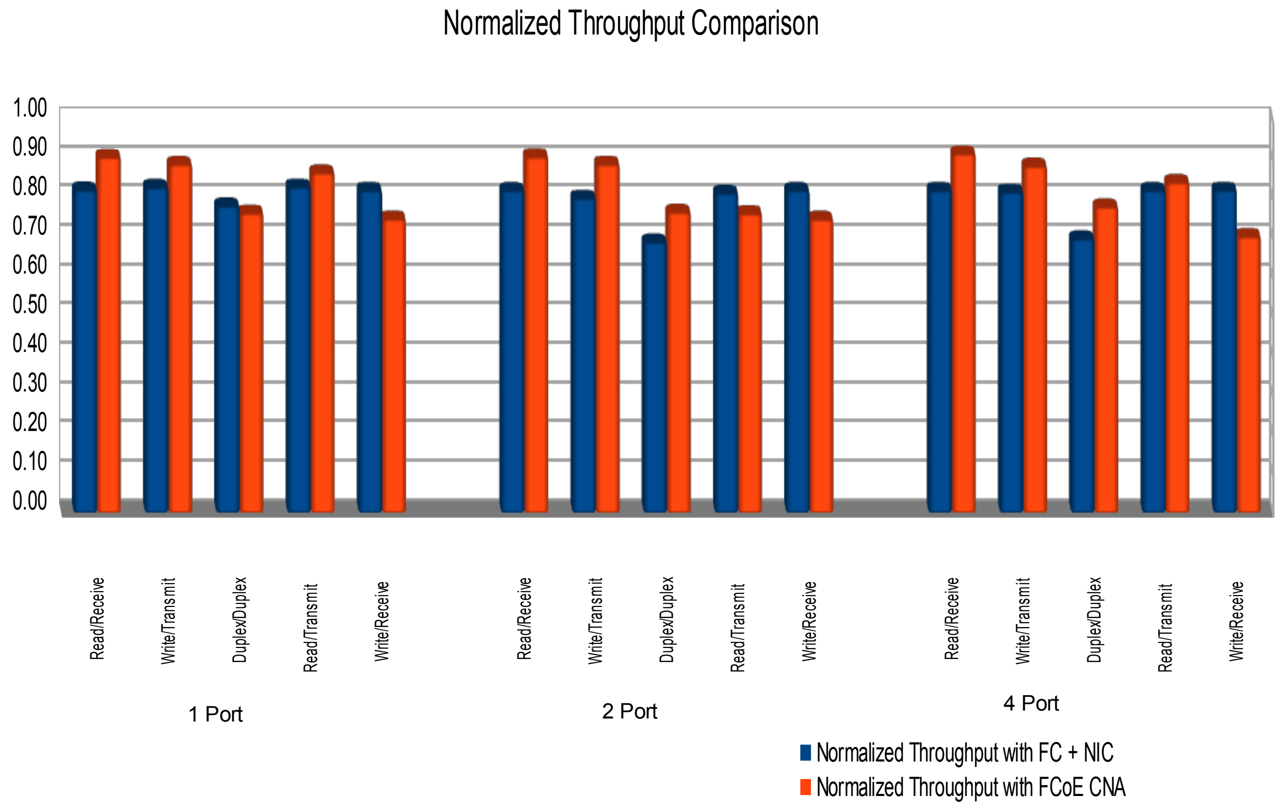

Fig 5 : Normalized Throughput Comparison

Another related observation is shown in the chart below. We have made a comparison using 2 CN4058 ports in the place of 1 FC3172 port and 1 EN4054 Ethernet port, in order to showcase the throughput improvement for combined Ethernet and FC traffic achieved using the same number of ports.

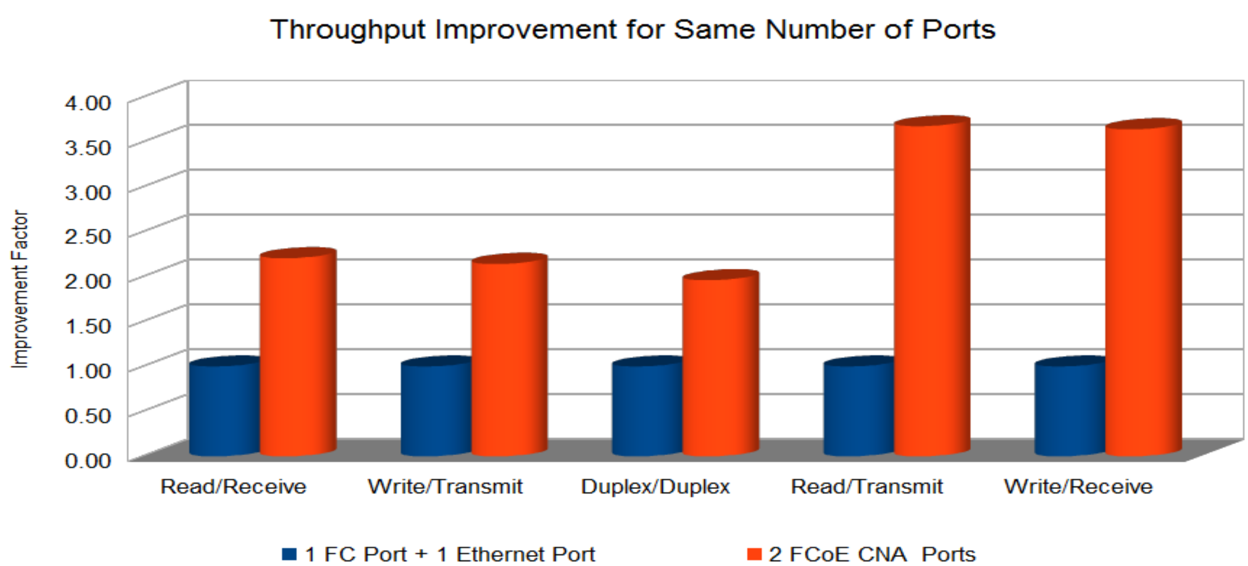

Fig 6: Throughput Improvement for Same Number of Ports

We find that there is at least 2x performance improvement with FCoE CNA. More importantly, we see almost 3.5x performance improvement in the Read/Transmit and Write/Receive patterns. This improvement happens as a result of the complete utilization of all the channels in the CNA during such scenarios. In contrast, when a FC port and Ethernet port are used, only one channel of each port is utilized. This means that customer workloads with such patterns can leverage the throughput advantage provided by the FCoE CNA.

Further, the total number of PCI slots, adapters, ports, cabling and switches required to achieve the same throughput using FCoE CNA is less compared to the traditional data center setup. The table illustrated below shows the CAPEX savings on the total number of components required for a typical data center with 26 servers.

Components	Ethernet	FC	Traditional Setup (FC+Ethernet)	CNA	FCoE Setup	Savings
Adapters	26	26	**52**	26	**26**	**50.00%**
Switches	2	2	**4**	2	**2**	**50.00%**
Cables	56	56	**112**	60	**60**	**46.00%**

Fig 7: Hardware Savings Comparison

There are various implications of having reduced number of components in the data center. Lesser the number of hardware components, lesser the power consumption in the FCoE environment, which would mean huge savings in power. The fact that there are still open adapter slots and open switch slots means there is the possibility of either scaling out or providing some other protocol via those slots. Also there is a cost saving attached when there are lesser hardware components.

5.2 Throughput Scaling for Multiple CNA Ports

In this section we demonstrate the throughput scalability of FCoE CNA with increase in the number of ports. This means that as customer requirements scale, the infrastructure is able to scale and meet the throughput requirements.

The below chart shows the measured total throughput for FCoE and Ethernet traffic for 1, 2, 4 and 8 ports of CN4058 CNA .

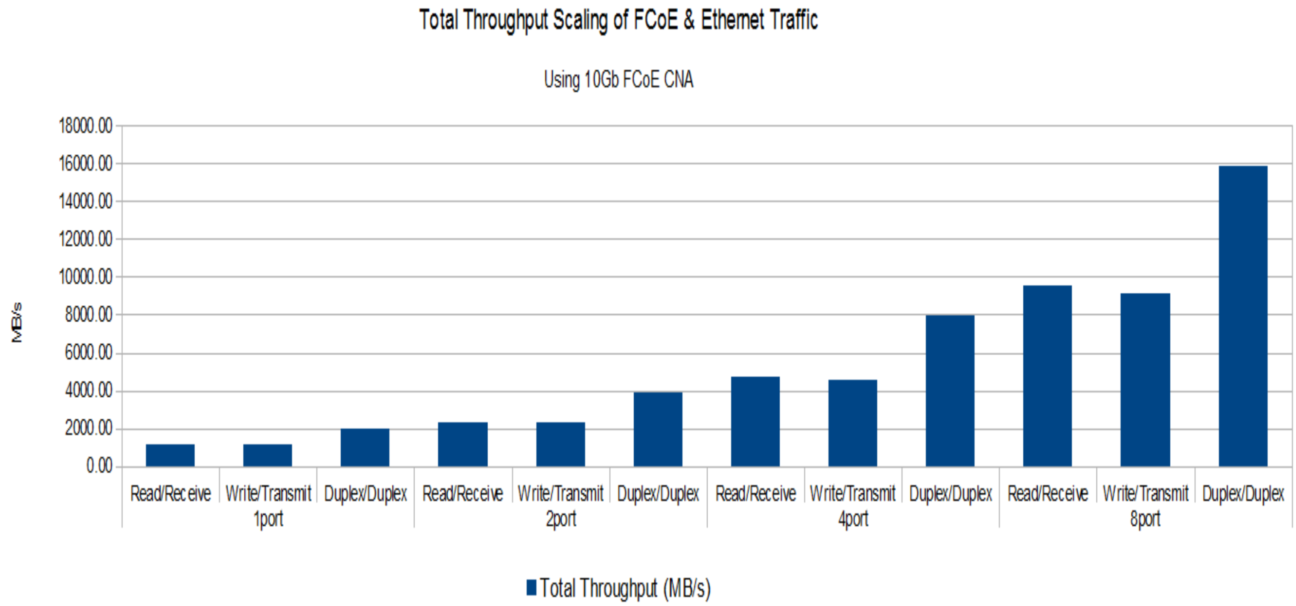

Fig 8: Total Throughput Scaling of FCoE and Ethernet Traffic

5.3 Bandwidth allocation for FCoE and Ethernet traffic using a FCoE CNA

As mentioned in section 1.2, Enhanced Transmission Selection (ETS) provides a method for allocating port bandwidth based on 802.1p priority values in the VLAN tag. ETS is defined in IEEE 802.1Qaz. Using ETS, different amounts of link bandwidth can be specified for different traffic types (such as for LAN, SAN, and management). ETS is an essential component in a converged environment that carries different types of traffic, each of which is sensitive to different handling criteria, such as Storage Area Networks (SANs) that are sensitive to packet loss, and LAN applications that may be latency-sensitive. In a single converged link, such as when implementing FCoE, ETS allows SAN and LAN traffic to coexist without imposing contrary handling requirements upon each other.

In our test setup, the converged switch provides the ETS functionality, thereby allowing the bandwidth allocation for SAN and LAN traffic. Here we discuss two scenarios, where the bandwidth allocation for SAN and LAN was varied and the throughput was measured.

Case 1: FC & Ethernet Bandwidth Proportion With 10(LAN)/50(SAN)/40 Bandwidth Allocation

In this scenario, the LAN traffic was allocated 10% of the total bandwidth and the SAN 50%. 40% was allocated for any other latency sensitive traffic. In our environment, since we used only network and storage workload, the 40% allocation for other traffic would be consumed between these workloads.

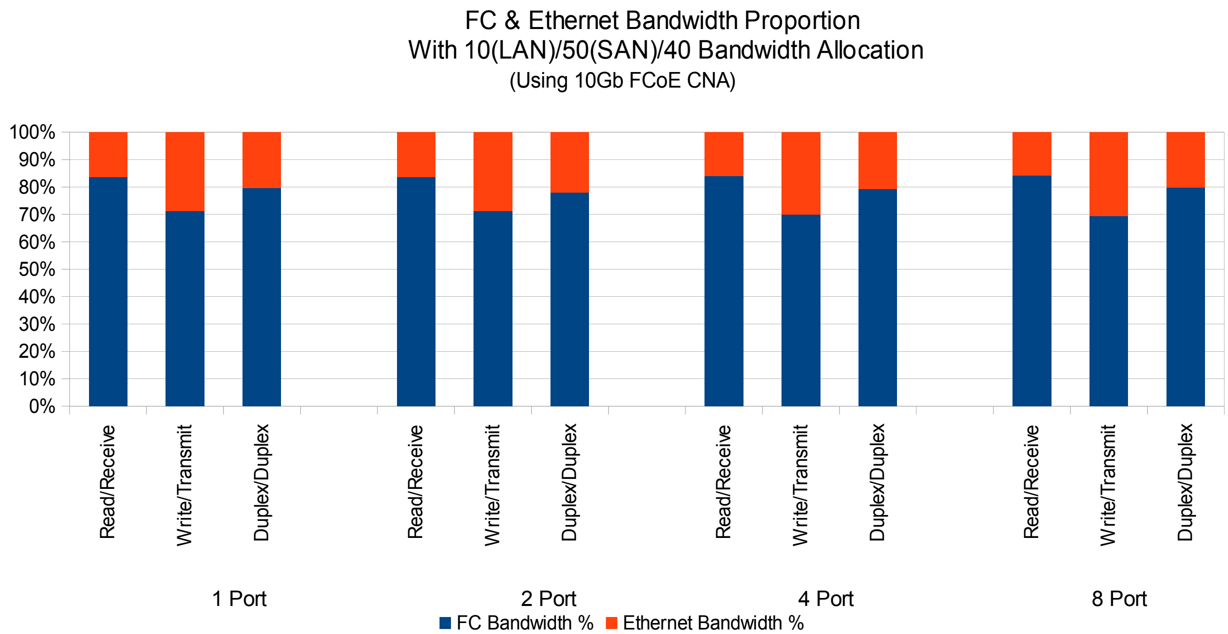

Fig 9: FCoE and Ethernet Bandwidth Proportionality for Case 1

Case 2: FC & Ethernet Bandwidth Proportion With 20(LAN)/70(SAN)/10 Bandwidth Allocation

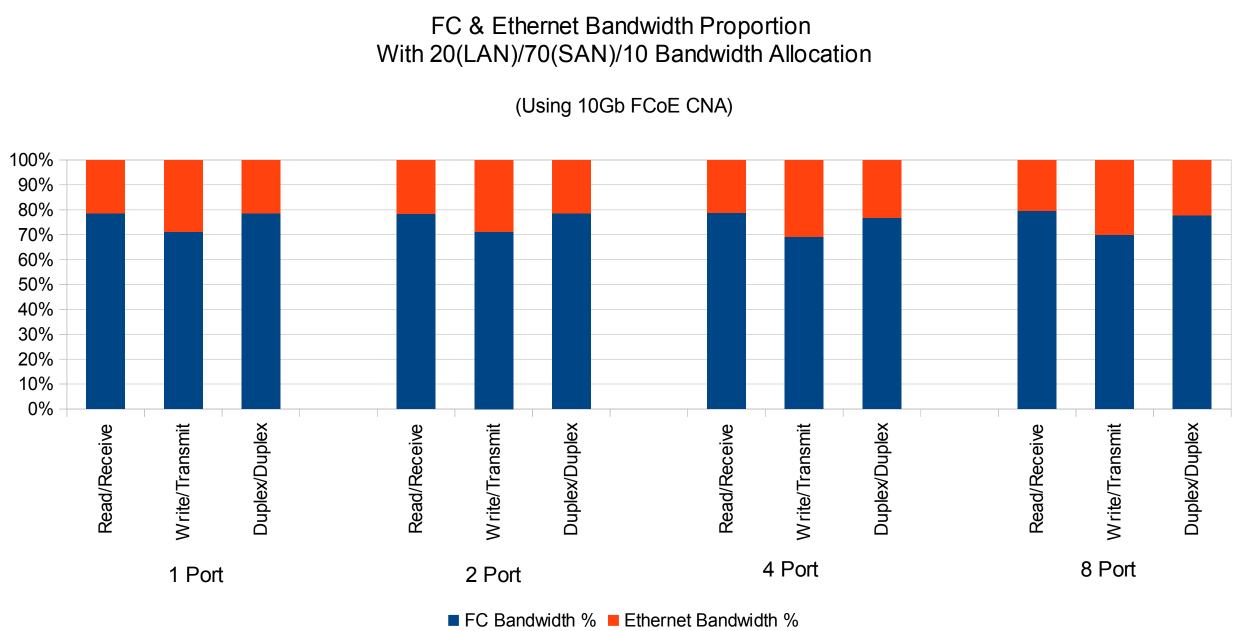

Fig 10: FCoE and Ethernet Bandwidth Proportionality for Case 2

In this scenario, the LAN traffic was allocated 20% of the total bandwidth and the SAN 70%. 10% was allocated for any other latency sensitive traffic. In our environment, since we used only network and storage workload, the 10% allocation for other traffic would be consumed between these workloads.

In both the above bandwidth proportionality experiments, we found that the FCoE CNA and the converged switch helps to manage the bandwidth proportions between the two types of traffic.

6. Conclusion

In the current virtualized data centers multi-core CPUs need to allow multiple workloads to be consolidated on the same server. In order to achieve that, requirements like IO consolidation is essential. Customers find a need in using reduced hardware components, thereby reducing their power and cooling cost. There is also a drive towards having simplified management and dynamic provisioning tools and capabilities.

Given these trends and requirements, storage and networking convergence through FCoE Converged Network Adapters have a major role to play in the current data centers. A CNA provides superior IO performance, increases server consolidation and frees up CPU from protocol processing. In this paper, we showcased the performance comparison of FCoE CNA over a traditional setup with separate FC and NIC adapters. We also discussed the throughput scalability obtained using FCoE CNA and how features like ETS can benefit in a converged environment with different types of traffic. More importantly, we demonstrated that replacing a traditional setup of separate FC and Ethernet ports with the same number of FCoE CNA ports, we could get almost 2x higher performance thereby providing savings on cost, power and space.

Hence this paper provides strong evidence that would encourage our customers to adopt CNA technology and integrate it into their data center infrastructure.

7. Acknowledgements

We would like to thank Jason P Kolhoff for helping us with all the infrastructure setup required for collecting the data. We also thank Clark Anderson, Ingo Dimmer and Bernard King-Smith in reviewing and helping us improve the whitepaper.

Disclaimer : The views and opinions expressed in this white paper are purely that of the authors and not necessarily that of IBM.