NETWORK ATTACHED STORAGE INTEROPERABILITY TESTING

Interoperability Test Results for Juniper Networks EX Series Ethernet Switches and NetApp Storage Systems

Advantage Alliance Partner

INTEGRATED ARCHIVE SYSTEMS

Table of Contents

Table of Figures

Introduction

Which technology is better for today's storage environment: the traditional storage area network (SAN) utilizing Fibre Channel interconnects, or IP-based storage protocols that use Ethernet and TCP/IP connections?

This question has long been debated by the storage and networking industries. SAN proponents argue that Fibre Channel provides several features—such as a lossless fabric, in-order delivery, and remote state change notification, among others—that are critical in storage environments. Ethernet and TCP/IP networks, by contrast, are lossy, subject to congestion, and impose a significantly higher overhead on the host and storage systems.

While these limitations have for some storage vendors precluded Ethernet and TCP/IP from consideration as a reliable medium for storage traffic, the emergence of low latency, line-rate Ethernet switches such as the Juniper Networks® EX Series Ethernet Switches, exponential growth in server and storage system processing power, and the advent of TCP Offload Engine (TOE) network interface cards have all but eliminated these concerns. Not only that, but Gigabit Ethernet (GbE) and 10GbE solutions offer a significantly lower cost per port compared to Fibre Channel SANs. Therefore, it is no surprise that enterprises of all sizes are deploying business critical applications such as Oracle databases and VMware Virtual Infrastructure/vSphere using IP-based storage protocols over GbE and 10GbE.

Scope

This application note is designed to demonstrate that the EX Series Ethernet switches from Juniper Networks interoperate with IP storage protocols such as the Internet Small Computer System Interface (iSCSI), the Common Internet File System (CIFS), and the Network File System (NFS), offering a high-performance storage switching solution.

The tests described in this document were jointly sponsored by Juniper Networks and NetApp and conducted by Integrated Archive Systems (IAS), a Juniper/NetApp partner. They indicate that EX Series switches, working with NetApp FAS3040c storage systems, deliver an effective and interoperable storage networking solution. Because of their superior capabilities, users can confidently deploy EX Series switches as part of their IP storage solution without fear of sacrificing efficiency, interoperability, or performance.

Design Considerations

Juniper Networks EX Series Ethernet Switches

The EX Series switches—specifically the Juniper Networks EX4200 Ethernet Switch with Virtual Chassis technology and the EX8208 modular Ethernet switch—provide high-performance, scalable solutions for data center core, top-of-rack, and end-of-row applications.

EX4200 switches deliver the high availability and high port densities of a modular chassis in a compact, cost-effective, pay-as-you-grow platform, delivering a scalable solution for today's evolving data center. Featuring Juniper's unique Virtual Chassis technology, the EX4200 switches let users start with a single switch and easily add more as business demands dictate. Up to 10 interconnected switches can operate as a single logical device, sharing a single operating system and configuration file. As a result, when configured as a Virtual Chassis, the EX4200 Ethernet Switch lowers total cost of ownership by reducing the number of managed devices by up to a factor of 10.

The EX8208 modular Ethernet switch offers a high-density, high-performance, and highly available platform for the most demanding data center core, cloud computing, and enterprise campus environments, as well as for aggregating top-of-rack or end-of-row access switches in the data center. The EX8208 can accommodate a variety of wire-rate 48-port Gigabit Ethernet (GbE) and 10GbE line cards, in any combination, to support existing requirements. Fully equipped, the eight slot EX8208 can support 384 GbE ports or 64 10GbE ports.

NetApp Storage Systems

The NetApp FAS3000 Series is a unified storage system providing high-performance SAN and network-attached storage (NAS) functionality. NetApp's FAS series storage systems provide industry leading integration with business applications that are prevalent in today's enterprises such as Oracle, SAP, VMware, Microsoft SQL Server, Microsoft Exchange, and Microsoft SharePoint.

NetApp's robust unified storage architecture, combined with its powerful data management software, provides flexibility, scalability, performance, optimized capacity utilization, and consolidated capacity and data protection management.

Test Summary

Three primary IP storage protocols are in wide use today: iSCSI, CIFS, and NFS.

For this application note, all three protocols were tested on the EX Series switches using two different operating systems: Windows (for iSCSI and CIFS) and Linux (for NFS). A variety of storage I/O tests, in which Ethernet storage traffic was sent over both 1 gigabit per second (Gbps) and 10 Gbps interfaces, were conducted to validate the performance of each protocol over Ethernet running through the EX Series switches.

Each test was conducted twice: once with three EX4200 switches deployed in a Virtual Chassis configuration, and once with an EX8208 modular Ethernet switch. See Table 1 for a summary of the tests.

Table 1: Summary of Storage Protocol Tests on EX Series Switches

	EX4200	EX8208
Windows		
GbE iSCSI	✓	✓
10GbE iSCSI	✓	✓
GbE CIFS	✓	✓
10GbE CIFS	✓	✓
GbE iSCSI, 100% write	✓	✓
10GbE iSCSI, 100% write	✓	✓
GbE CIFS, 100% write	✓	✓
10GbE CIFS, 100% write	✓	✓
Linux		
GbE NFS	✓	✓
10GbE NFS	✓	✓
10GbE NFS with Routing Engine failover	✓	Not Tested
GbE NFS, 100% write	✓	✓
10GbE NFS, 100% write	✓	✓

Test Equipment

The following equipment was used to conduct the performance tests.

- EX4200 Ethernet switches: Total of three, deployed in a Virtual Chassis configuration
- EX8208 modular Ethernet switch: Configured with one EX8200-48T 10/100/1000BASE-T line card and two EX8200-8XS 10GbE line cards
- NetApp FAS3040c storage system
- Two Sun Microsystems Sun Fire X4150 servers running Windows Server 2003 SP2 or Red Hat Enterprise Linux (RHEL) 5 server
- Brocade 200e Fibre Channel Switch (for baseline performance)
- NetApp Simulated I/O (SIO) software

Test Setup and Background

During initial testing, successive runs were performed using block sizes ranging from 4,000 to 256,000, working sets ranging from 5 MB to 2 GB, and thread counts ranging from 8 to 32. Because the NetApp FAS storage controller handles read and write I/O requests differently, disparate working sets were used for the read and write testing scenarios.
The best throughput for all "read" pretests was achieved using 28 to 32 threads on both Windows 2003 and Red Hat Enterprise Linux. The final setting for all test scenarios was 28 threads, which in some cases produced marginally better throughput than 32 threads.

For all “write” pretests, the best throughput was achieved using eight threads on both Windows 2003 and Red Hat Enterprise Linux. All tests were run for 60 seconds to ensure a sufficient sampling of peaks and an adequate statistical mean.

The “read” tests were conducted with a large sequential I/O profile using a very small data set. For these scenarios, the SIO parameters were set to 100% reads using a 256,000 block size with 0% randomness against a 5 MB working set. The 5 MB working set assured that the data would be served directly from cache, eliminating the disk as a potential bottleneck. In the performance test, a 256,000 block size produced the highest throughput; all of the NFS “read” tests used the direct switch for storage I/O to ensure that all data was served directly from the storage controller and not from the client’s local cache.

Although the “write” scenarios were also conducted with a large sequential I/O profile, the working set size changed significantly. For these test scenarios, the storage I/O parameters were set to 100% writes using a 256,000 block size with 0% randomness against a 512 MB working set. A 2 GB file was used as the target; however, a 512 MB working set produced the best results. Theoretically, an 88,000 working set should provide the best write throughput since, on average, only one block would be placed on each data disk. However, a larger working set ensured a more statistically even distribution of blocks and I/O requests across all disks in the aggregate. In pretesting, a 256,000 block provided the optimal throughput.

Rather than use a Link Aggregation Control Protocol (LACP) link aggregation group on the NetApp controller, two separate GbE connections were used for the iSCSI GbE testing scenarios to comply with NetApp best practices (see the NetApp Technical Report *iSCSI Multipathing Possibilities on Windows with Data ONTAP* at **http://media.netapp.com/documents/tr-3441.pdf**). The NetApp device used in the GbE NFS testing used an LACP-type dynamic multimode virtual interface (vif)[1], including two GbE links. In this case, two of the onboard GbE links on the NetApp controller were used in the vif, with the clients each using a single GbE link for testing data traffic.

Baseline

To establish a performance baseline against which the EX Series switches could be compared, a set of tests were run with both Windows Server 2003 SP2 and Red Hat Enterprise Linux (RHEL) 5 using the Fibre Channel protocol. A Brocade 825 8Gbps Fibre Channel host bus adapter (HBA) was installed in each of the servers and connected to a Brocade 200e Fibre Channel switch. The client ports were then zoned to a single NetApp 4 Gbps target port.

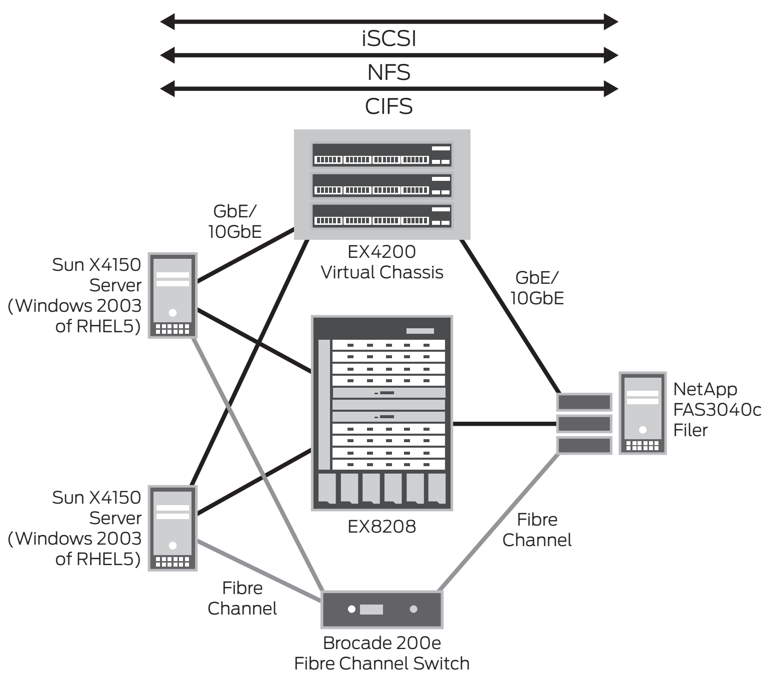

Figure 1: Storage over Ethernet test setup using EX Series switches.

[1] See Layton, Trey (2009). Multimode VIF survival guide. Ethernet Storage Guy Blog, Retrieved from http://communities.netapp.com/blogs/ethernetstorageguy/2009/04/04/multimode-vif-survival-guide

Test Results

Windows Tests

For the Windows environment, iSCSI and CIFS were tested through the EX4200 Virtual Chassis switches and EX8208 modular switches at 1 Gbps and 10 Gbps (see Figure 1). Peak and average results were measured in megabytes per second (MB/s) and gigabytes per second (GB/s). Both read and write traffic flows were measured for all test scenarios. See Table 2 for the results.

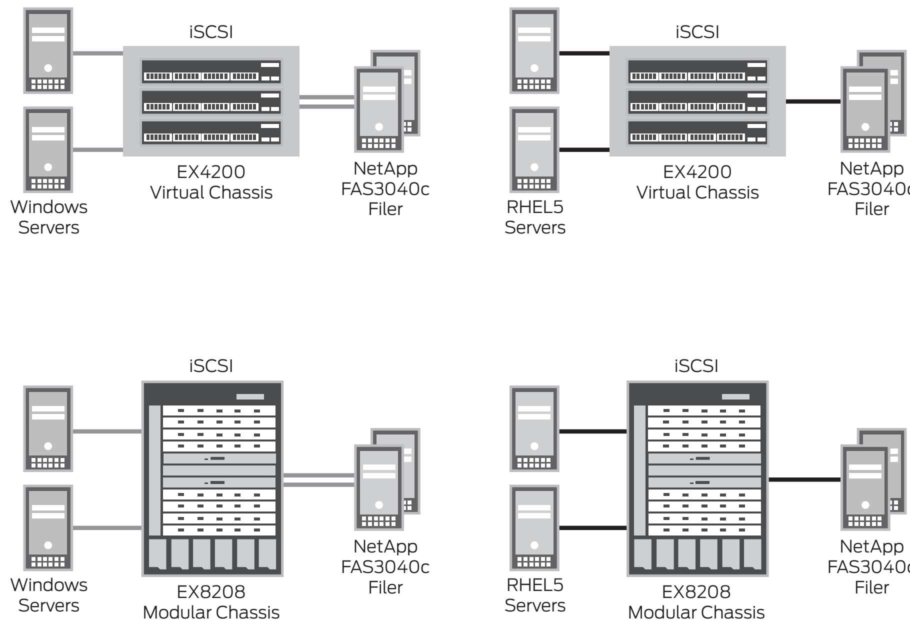

Figure 2: GbE and 10GbE iSCSI test configurations

Table 2: Windows Results for iSCSI and CIFS Through EX Series Switches with Fiber Channel Baseline

	PEAK THROUGHPUT (MB/S)	PEAK THROUGHPUT (GB/S)	60 SECOND AVERAGE THROUGHPUT (MB/S)	60 SECOND AVERAGE THROUGHPUT (GB/S)
EX4200				
GbE iSCSI, 100% read	239.23	1.91	224.47	1.80
10GbE iSCSI, 100% read	525.67	4.21	508.36	4.07
GbE CIFS, 100% read	236.93	1.90	222.13	1.78
10GbE CIFS, 100% read	518.30	4.15	497.53	3.98
GbE iSCSI, 100% write	189.77	1.52	157.42	1.26
10GbE iSCSI, 100% write	283.93	2.27	231.88	1.86
GbE CIFS, 100% write	157.08	1.26	130.48	1.04
10GbE CIFS, 100% write	284.34	2.27	240.17	1.92
EX8208				
GbE iSCSI, 100% read	231.21	1.85	225.08	1.80
10GbE iSCSI, 100% read	493.34	3.95	478.38	3.83
GbE CIFS, 100% read	220.91	1.77	213.50	1.71
10GbE CIFS, 100% read	479.30	3.83	462.04	3.70
GbE iSCSI, 100% write	206.15	1.65	178.72	1.43
10GbE iSCSI, 100% write	272.19	2.18	242.88	1.94
GbE CIFS, 100% write	210.78	1.69	184.11	1.47
10GbE CIFS, 100% write	273.86	2.19	241.58	1.93
Baseline: 4 Gb Windows Fibre Channel 100% read	198.14	1.59	192.20	1.54

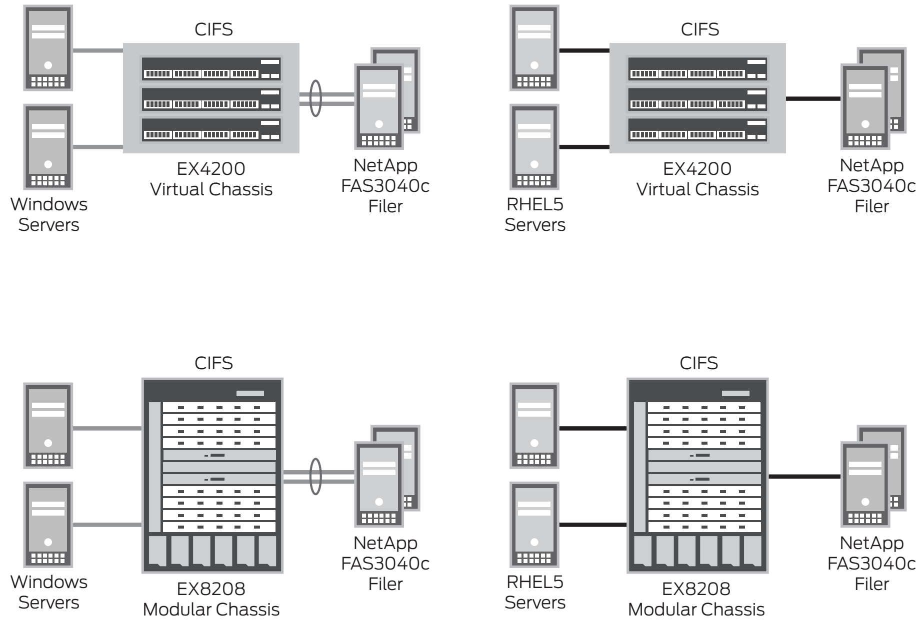

Figure 3: GbE and 10GbE CIFS test configurations

Linux Tests

For the Linux environment, NFS was tested through the EX4200 Virtual Chassis switches and EX8208 modular switches at 1 Gbps and 10 Gbps (see Figure 2). Storage controller peak and server I/O aggregate results were measured in Mbps and Gbps. Both read and write traffic flows were measured for all test scenarios. See Table 3 for the results.

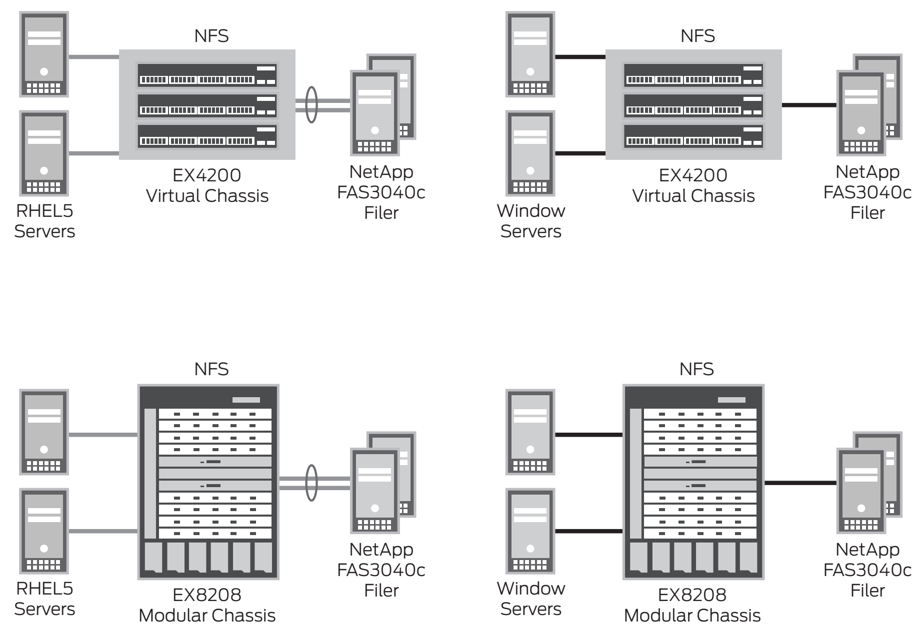

Figure 4: GbE and 10GbE NFS test configurations

Linux Results for NFS Through EX Series Switches with Fibre Channel Baseline.

Table 3: Linux Results for NFS Through EX Series Switches with Fibre Channel Baseline

	PEAK THROUGHPUT (MBPS)	PEAK THROUGHPUT (GBPS)	60 SECOND AVERAGE THROUGHPUT (MBPS)	60 SECOND AVERAGE THROUGHPUT (GBPS)
EX4200				
GbE NFS, 100% read	240.60	1.92	223.82	1.79
10GbE NFS, 100% read	580.50	4.64	512.57	4.10
10GbE NFS with mid-test Routing Engine failover	586.53	4.69	499.57	4.00
GbE NFS, 100% write	121.90	0.98	64.01	0.51
10GbE NFS, 100% write	263.67	2.11	221.56	1.77
EX8208				
GbE NFS	240.46	1.92	223.82	1.79
10GbE NFS	475.80	3.81	406.13	3.25
GbE NFS, 100% write	247.88	1.98	222.72	1.78
10GbE NFS, 100% write	475.80	3.81	362.61	2.90
Baseline: 4 Gb Linux Fibre Channel 100% read	174.53	1.40	156.61	1.25

10-Gigabit Ethernet (10GbE) Results

Peak performance for the 10GbE 100% read tests on the EX8208 ranged from 3.81 to 3.95 Gbps, with average performance ranging from 3.25 to 3.83 Gbps. For the 100% write tests, peak performance ranged from 2.18 to 2.95 Gbps, with the average ranging from 1.93 to 2.34 Gbps.

On the EX4200 switches, peak performance for the 10GbE 100% read tests ranged from 4.15 to 4.69 Gbps, averaging between 3.98 and 4.1 Gbps. On the 100% write tests, peak performance ranged from 2.11 to 2.27 Gbps, averaging between 1.77 and 1.92 Gbps.

For the read tests, the NetApp device's sysstat command indicated that 100% of the data was being served from cache; however, even though CPU utilization was typically between 15-25%, the system could not be pushed past 4.69 Gbps. On the write tests, performance was limited by the number of disks that were available for the test. The output of the NetApp sysstat command during the write tests shows that the system was running at 80-90% CPU utilization.

GbE Results

Peak performance for the GbE 100% read tests on the EX8208 ranged from 1.77 to 1.92 Gbps, with an average performance ranging from 1.71 to 1.80 Gbps. For the 100% write tests, peak performance ranged from 1.65 to 1.98 Gbps, with the average ranging from 1.43 to 1.78 Gbps.

On the EX4200 switches, peak performance for the 100% read tests ranged from 1.9 to 1.92 Gbps, with an average of between 1.78 and 1.8 Gbps. For the 100% write tests, peak performance ranged from .98 to 1.52 Gbps, with an average of .51 to 1.26 Gbps.

For the GbE tests, the bottleneck appeared to be limited bandwidth. As the test results show, the clients were saturating the two GbE network connections on the NetApp controller. For the read tests, the cache hit rate remained at 100% and CPU utilization remained between 15-25%. On the write tests, the NetApp device spent between 98-100% in a check point state, while CPU utilization was between 67-90%.

Summary: EX Series Delivers Credible Ethernet Storage Solution

In the read tests, the EX4200 and EX8208 switches posted very favorable results using Ethernet storage protocols as compared to the 4 Gbps Fibre Channel baseline, with 141-236% better throughput over 10GbE, and 11-37% better throughput over GbE. These results, combined with low latency, line-rate performance, deep per-port buffers, and the tested and proven Juniper Networks Junos® operating system, demonstrate that EX4200 and EX8200 switches in combination with Ethernet storage protocols deliver a compelling and credible storage platform.

Appendix: Storage Protocols

CIFS: The Common Internet File System (CIFS) is an IP-based protocol that defines a standard for remote file access, enabling users with different platforms and computers to share files without having to install new software. Although CIFS runs over TCP/IP, it uses the Server Message Block (SMB) protocol found in Microsoft Windows for file and printer access, enabling any application to open and share files across the Internet.

iSCSI: An abbreviation of Internet Small Computer System Interface, iSCSI is an IP-based standard for linking data storage facilities, facilitating data transfers over intranets and managing storage over long distances. The iSCSI standard can be used to transmit data over LANs and WANs as well as the Internet, and can enable data storage and retrieval that is independent of location.

NFS: Developed by Sun Microsystems, the Network File System (NFS) file-sharing protocol is a de facto Unix standard widely known as a "distributed file system." The name is somewhat misleading, though, because NFS is not a disk file system that reads and writes the disk sectors but instead enables the operating system to view files on networked computers as if they were local.

About Juniper Networks

Juniper Networks, Inc. is the leader in high-performance networking. Juniper offers a high-performance network infrastructure that creates a responsive and trusted environment for accelerating the deployment of services and applications over a single network. This fuels high-performance businesses. Additional information can be found at **www.juniper.net**.

Corporate and Sales Headquarters
Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089 USA
Phone: 888.JUNIPER (888.586.4737)
or 408.745.2000
Fax: 408.745.2100
www.juniper.net

APAC Headquarters
Juniper Networks (Hong Kong)
26/F, Cityplaza One
1111 King's Road
Taikoo Shing, Hong Kong
Phone: 852.2332.3636
Fax: 852.2574.7803

EMEA Headquarters
Juniper Networks Ireland
Airside Business Park
Swords, County Dublin, Ireland
Phone: 35.31.8903.600
EMEA Sales: 00800.4586.4737
Fax: 35.31.8903.601

3500171-003-EN Jan 2012

Printed on recycled paper