THE TOLLY GROUP

No. 208351 March 2009

BlueArc Corporation

Titan 3210 Network Storage System

Throughput Evaluation with NFS Version 3 Protocol

TOLLY Up to Spec CERTIFIED

Test Summary

Premise: *As high performance applications and computing environments proliferate, enterprise network managers demand that their network storage systems offer high I/O performance, scalability and expandability to support ever-evolving, mission-critical applications and protect their investment. Understanding true performance characteristics provided by each storage system in various scenarios, including clustering, is essential for network managers to make an informed buying decision.*

BlueArc Corporation commissioned The Tolly Group to examine the I/O performance characteristics of the Titan 3210 Network Storage System, a purpose-built network storage solution.

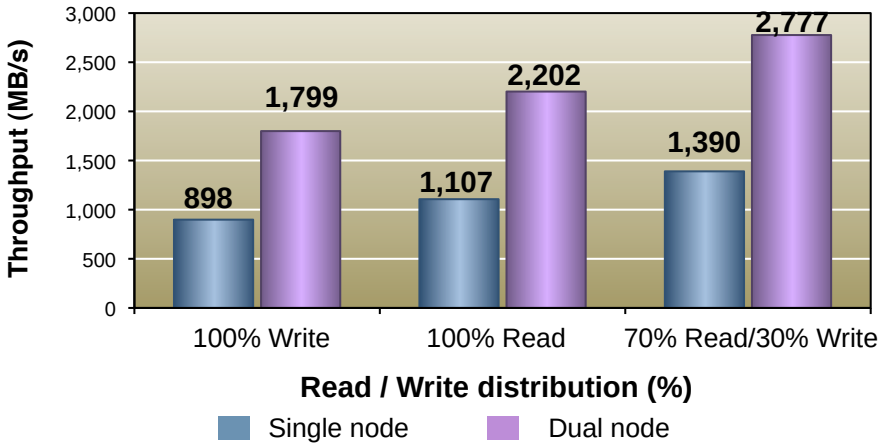
Tolly Group engineers measured the throughput of a single Titan 3210 node when handling file-access requests using the Network File Systems (NFS) Version 3 protocol. The back-end storage system used for testing was a Hitachi Data Systems (HDS) Universal Storage Platform V (USP V) outfitted with 256 Fibre Channel (FC) HDDs. Throughput was measured for a variety of read, write and mixed read/write operations from/to back-end disks with sequential access. Engineers also measured the throughput of two Titan 3210 nodes in a cluster in the same scenarios.

Tests were conducted in October 2008. (BlueArc and HDS are strategic partners of high-performance storage solutions.)

Test Highlights

- Delivers near wire-speed throughput when handling 100% sequential read operations over 10GbE connections
- Provides linear scalability by achieving twice as much throughput with two nodes as a single node for all tests
- Sustains average throughput of 891 MB/s in a single node and 1,785 MB/s for dual-node clusters in tests of 100% sequential writes for 32KB, 64KB and 128KB block sizes
- Proves it can offer optimized, scalable high-performance NAS services for the NFS Ver. 3 environment running I/O intensive operations

NFS Ver. 3 Throughput for Various I/O Operations Using 32KB Blocks in Single- and Dual-Node Cluster Scenarios
as Reported by IOzone 3.283



Note: Each of the six clients generated eight concurrent workload requests, resulting in 48 total threads. Each thread read 5 GB of data, resulting in a total of 240 GB of data retrieved via a single BlueArc Titan 3210 using 100% sequential disk access and 1,500-byte MTU. The above remains the same for all the test scenarios.

Source: The Tolly Group, October 2008

Figure 1

Executive Summary

BlueArc's Titan 3210 Network Storage System delivers near wire-speed, 10GbE throughput conducting sequential read operations for various block sizes using NFS file system and doubles its performance in the same tests by clustering two nodes.

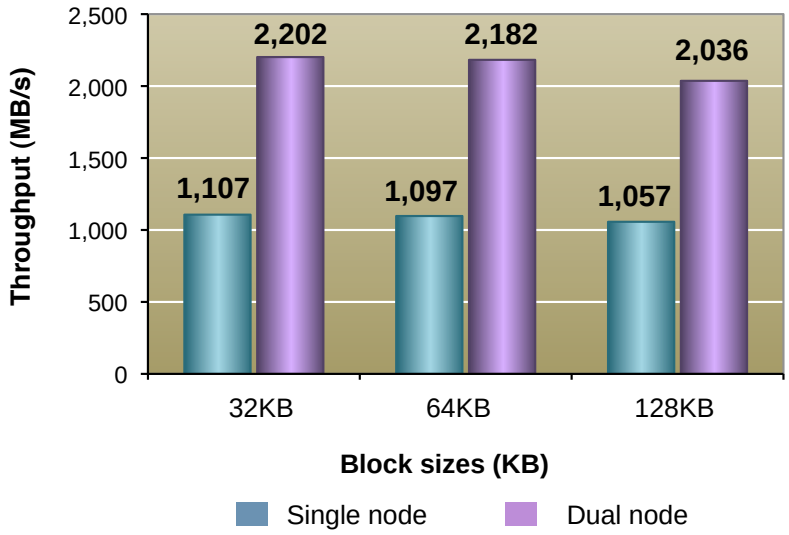
With the rapid increase in the amount of data handled by today's enterprise infrastructures, faster and more scalable storage systems are required to keep up.

Conventional, general-purpose network-attached storage (NAS) systems leverage CPU-based architectures resulting in serial processing. As performance demands increase, this requires the deployment of additional filters to meet workloads. As such, conventional NAS system are not optimized for performance-oriented file-system processing and disk access.

Conversely, BlueArc's hardware-based architecture leverages field programmable gate arrays (FPGAs) to perform operations in parallel. This facilitates high throughput between servers and a dedicated storage backend with massive scalability.

This test focused on measuring throughput of a BlueArc Titan 3210 hardware-based network storage solution that accessed the NFS file system using various block sizes in a mix of read/write operations, and comparing

NFS Ver. 3 100% Sequential Read Throughput for Various Block Sizes in Single- and Dual-Node Clusters as Reported by IOzone 3.283



Source: The Tolly Group, October 2008 *Figure 2*

that with throughput from a dual-node cluster. NFS is one of two most commonly used file systems in NAS and the standard file sharing mechanism for Unix/Linux environments.

The Tolly Group's hands-on evaluation of the BlueArc Titan 3210 Network Storage System demonstrates that the Titan 3210 delivered 1,100 MB/s (close to 10 Gbps considering network overhead) of throughput when handling 100% sequential read operations using a single node and 2,200 MB/s (close to 20 Gbps) using a dual-node cluster. During 100% write operation, the Titan 3210 delivered about 900 MB/s and 1,800 MB/s with a single-node and dual-node cluster, respectively. During a mixed read/write test, throughput exceeded 10 Gbps since read and write operations use two separate 10-Gbps channels. Overall, throughput was consistent across three different block sizes tested.

The above results clearly show that a single Titan 3210 system offers excellent throughput using NFS as a file sharing mechanism in Unix/Linux environments. This also proves that its enhanced protocol acceleration helps to achieve the full potential of NFS performance by optimizing its file processing. When two Titan 3210 systems are clustered, the tandem delivers even more impressive performance by doubling the single-node performance without showing any sign of degradation, which is normally caused by the cluster overhead. Comparatively, traditional NAS systems exhibit difficulty in overcoming the cluster overhead issue. It is very common to set up more than three nodes to achieve twice the performance of the single node but adding more nodes into a cluster introduces another serious question about the reliability of these systems.

Results

100% Sequential Read Operation

Test results show that during 100% sequential read operations using one node, the Titan 3210 achieved I/O throughput of 1,107, 1,097 and 1,057 MB/s for 32KB,

64KB and 128KB block sizes, respectively. These convert to 8.9, 8.8 and 8.5 Gbps of throughput. These throughput values represent near wire-speed performance over a 10GbE connection. *(Note: default MTU of 1,500 bytes was used throughout the test.)*

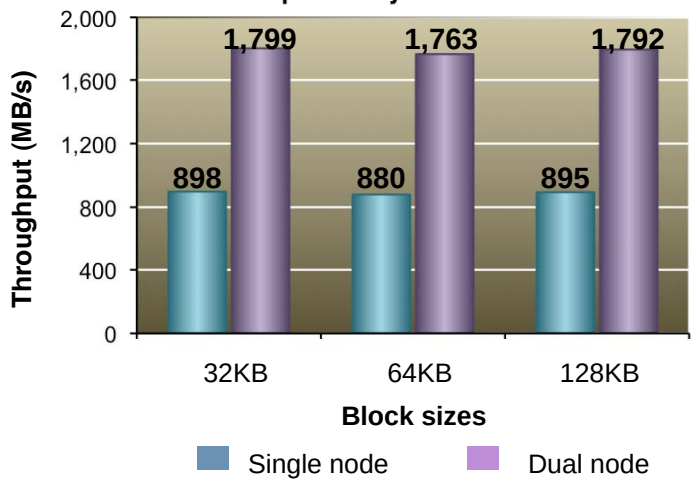
When clustering two nodes, engineers validated that throughput is simply doubled by measuring 1.99X, 1.99X and 1.93X, as much as the single node for 32KB, 64KB and 128KB block sizes. See Figure 2.

About caching...

While caching data will help performance and is recommended in real-world deployments, our focus in this test was to push actual storage systems to their limits by forcing data to be retrieved from the actual disks. This test also took advantage of caching but, engineers deliberately configured cache sizes to be relatively small or in its default configuration to keep the caching effect under nominal level. For this reason, engineers configured the clients with only 512 MB RAM to minimize the possible client-side caching and also limit the cache memory on the HDS USP V storage system to 32 GB per node even though the tested system was outfitted with 256 GB of cache memory. The default caching configuration of the Titan 3210 was used.

This effort cannot be complete without an appropriate choice of test data set. In the read test, each of the six clients generated eight concurrent workload requests, resulting in 48 total threads. Each thread read 5 GB of data from the back-end HDS USP V storage system, resulting in a total of 240 GB of data to be retrieved via a single BlueArc Titan 3210 and 480 GB via dual-node Titan 3210 cluster using 100% sequential disk access and 1,500-byte MTU. A portion of the data was served by the caches across multiple phases but the majority of data had to be retrieved from the disks in the backend storage system.

NFS Ver. 3 100% Sequential Write Throughput for Various Block Sizes in Single- and Dual-Node Clusters
as Reported by IOzone 3.283



Source: The Tolly Group, October 2008 *Figure 3*

BlueArc Corporation

Titan 3210 Network Storage System

Throughput Evaluation Using the Network File System (NFS) Protocol

Product Specifications

Vendor-supplied information not necessarily verified by The Tolly Group

BlueArc Titan 3210 Network Storage System

- Hardware-accelerated network storage with up to 20 Gbps throughput and up to 200,000 observed IOPs
- Dynamically scalable storage up to 4 PB under a single namespace, with file systems up to 256 TB
- Scalable N-Way high availability clustering technology
- Dynamic read caching for scalable read-intense workloads
- Supports over 16 million files per directory
- Unified NAS and SAN cluster namespace for unified directory structure
- Intelligent tiered storage across SSD, FC, SAS and SATA
- Virtual volumes and servers
- Integrated WORM file system
- Policy-based management & transparent data migration
- Advanced data protection and disaster recovery

For more information contact:
BlueArc Corporation
50 Rio Robles
San Jose, CA 95134, USA
Phone: 408-576-6600
Fax 408-576-6601
Sales Inquiries
sales@bluearc.com
866-864-1040

100% Sequential Write Operation

Test results show 898 MB/s, 880 MB/s, 895 MB/s throughput for 32KB, 64KB and 128KB block sizes. This demonstrates an approximate 15% to 19% decrease compared to the 100% read configuration.

Given that writing to disk is a more taxing I/O operation than reading from disk, especially when it writes large files (240 GB per node) at high speed, this level of 100% sequential write throughput is considered very good.

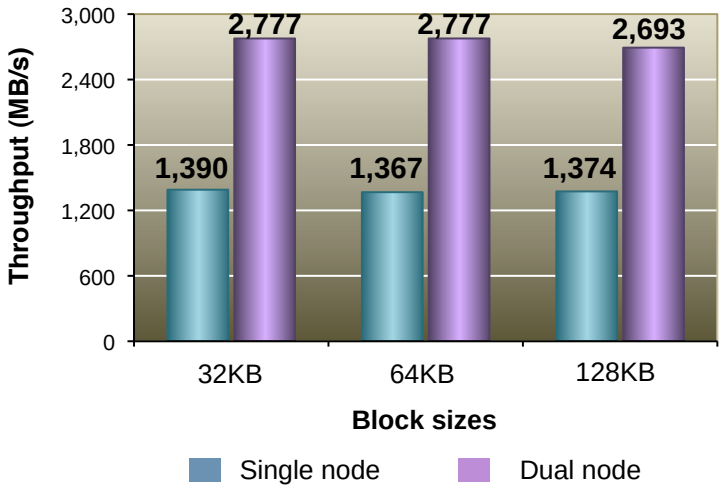
The Titan 3210's writing operation takes advantage of 4 GB of NVRAM inside one of the File System Modules (FSB). The NVRAM stores "writes" and returns fast acknowledgments to clients. This allows the clients to go on processing without waiting for the data to be written to disk. When NVRAM utilization reaches 50%, the Titan 3210 flushes any outstanding data to disk.

With a dual-node cluster configuration, the Titan 3210 reported throughput of 1,799, 1,763 and 1,792 MB/s for 32KB, 64KB and 128KB block sizes, which are is twice the single node performance. See Figure 3.

70% Read / 30% Write Operation

This test is unique in that read and write operations take place simultaneously. Because of that, this test can exercise the system's full-duplex capability. As expected, the aggregate I/O throughput exceeded the theoretical maximum of one 10GbE connection by measuring 1,390, 1,367 and 1,374 MB/s for 32KB, 64KB and 128KB block sizes in a single-node scenario. This is almost equivalent to 12 Gbps of throughput on the Ethernet physical layer. Again, the dual-node cluster results demonstrate twice the performance of the single node. See Figure 4.

NFS Ver. 3 70% Read / 30% Write Throughput for Various Block Sizes in Single-Node and Dual-Node Cluster Scenarios
as Reported by IOzone 3.283



Source: The Tolly Group, October 2008 *Figure 4*

Test Setup & Methodology

The test bed consisted of multiple components including a BlueArc Titan 3210 network storage system, clients, Force10 S2410P 10GbE LAN switch, SAN switch and HDS USP V storage system. See Figure 5 for detailed product information and Figure 7 for the test bed diagram.

The BlueArc Titan 3210 remained in its default configuration except for the NFS file system configuration for the testing. Engineers first created four storage pools per node and eight system drives (LUNs) were assigned to each pool. Second, engineers created one file system on each storage pool and then created one NFS share on each file system. With the configurations above, clients were able to mount the file shares using the NFS protocol.

A total of 12 client machines (six per node) were used for the test. Each client ran IOzone client program and connected to the IOzone control machine. IOzone Filesystem Benchmark tool (Ver. 3.283) was used for all tests. IOzone was run in throughput mode, which allowed all nodes to issue their I/Os simultaneously. A wrapper script was used to run IOzone iteratively with varying options.

For a single-node test, six client were used and IOzone was configured to use 32KB/64KB/128KB record sizes, for the 100% read, 100% write and 70% read/30% write tests. Each client generated eight threads, resulting in 48 threads, each reading/writing 5 GB of data. For the dual-node cluster, 12 client machines were used resulting in 96 threads. Everything else was the same as the single-node test.

A Brocade Silkworm 5320 was configured to connect between the Titan 3210's Fibre Channel ports and the HDS USP V's Fibre-

Detailed Component Summary of BlueArc High-performance Storage Solution Tested

Vendor	Model	Description	Version	Quantity
BlueArc Corporation	**Titan 3210**	**Network storage platform** with two 10 GbE ports, eight 4-Gbit/s Fibre Channel ports and 59 GB distributed memory	BOS 6.0	2
Hitachi Data Systems	**Universal Storage Platform V (USP V)**	**Large scale storage services platform**, equipped with 256GB cache memory (64GB cache memory or 32GB cache memory per node was used for this test), 16 Front-end Directors (FEDs) and eight Back-end Directors (BEDs). Each FED has eight 4-Gbit/s Fibre Channel ports. One Fibre Channel port per FED (total 8 FC ports per node) was used for this test. 256 FC HDDs (146 GB, 15K RPM) were configured for RAID-10 (64 X 2 + 2) and to create 64 LUNs (32 LUNS per node and 268 GB per LUN)	60-03-06-00	1
Sun Microsystems	**Sun Fire X2200 M2 Server**	2 X dual AMD Opteron processor 2218 HE 2.6 GHz, 512 MB RAM (used relatively small memory to minimize client-side caching) **OS:** Red Hat Enterprise Linux Server Release 5 64-bit edition (Kernel 2.6.18-8.e15 SMP) for NFS file service		12
Chelsio Communications	**S310E**	**10GbE Storage Accelerator** - a single 10 GbE port PCI Express 8x host bus interface with TCP offload. This is installed in each client. One per client.	1.0.146	12
Brocade Communications Systems	**Silkworm 5320**	**Fibre Channel (FC) switch** supporting 80 1, 2, 4 and 8 Gbit/s FC ports (Port speed was hardcoded to 4 Gbit/s for this test)	6.1.0a	1
Force10 Networks	**S2410P -- 24 XFP interfaces**	**Data Center Ethernet Switch** supporting 24 10 GbE ports	2.4.1.11	1

Source: The Tolly Group, October 2008 *Figure 5*

NFS Ver. 3 Results Summary: All Sequential Read/Write Scenarios Tested
as Reported by IOzone 3.283

Read / Write Distribution (%)	Block Size (KB)	I/O Throughput (MB/s)	
		One Node	Two-node Cluster
100% Write	**32 KB**	898	1,799
	64 KB	880	1,763
	128 KB	895	1,792
100% Read	**32 KB**	1,107	2,202
	64 KB	1,097	2,182
	128 KB	1,057	2,036
70% Read / 30% Write	**32 KB**	1,390	2,777
	64 KB	1,367	2,777
	128 KB	1,374	2,693

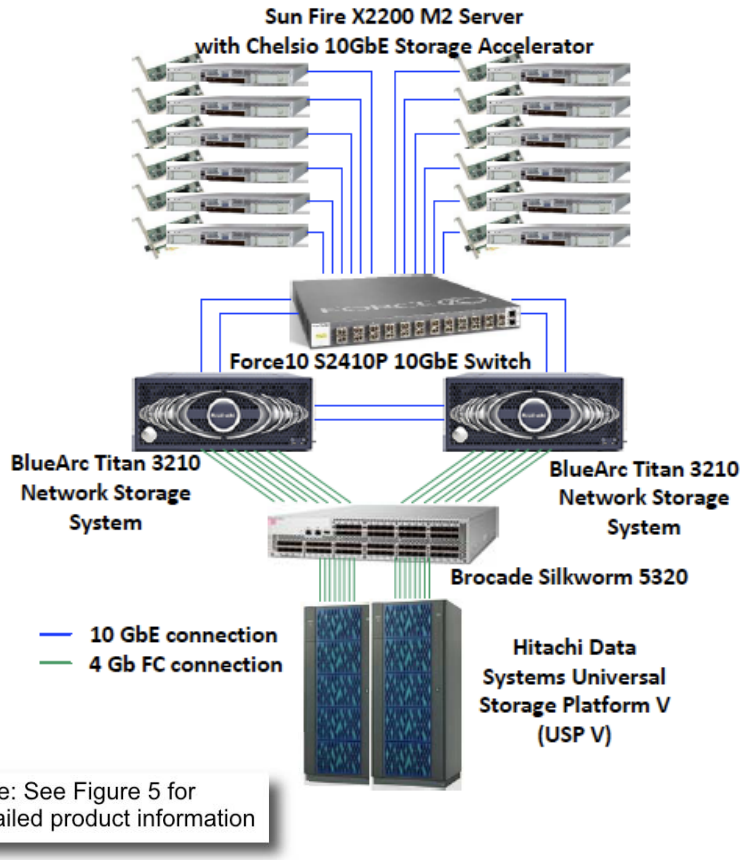
Source: The Tolly Group, October 2008 *Figure 6*

Channel ports. One-to-one connections were used and each pair was grouped into a separate zone, representing eight zones for the single-node test. However, there was only one zone configured for the two-node cluster test so that all 32 Fibre Channel ports were in the same zone.

HDS USP V was configured to use 32 GB of cache memory per node (64 GB for two nodes). Its 256 FC HDDs were configured for RAID 10 and 128 HDDs were supported for each node.

For network configuration, the Chelsio S310 10GbE storage adapter utilized its offload capabilities and standard Ethernet MTU (1,500 bytes). No Jumbo Frames were used for this test. A Force10 Networks 2410P 10GbE switch with 24 XFP interfaces remained in its default configura

Test Bed Diagram
BlueArc High-performance Storage Solution



Note: See Figure 5 for detailed product information

Source: The Tolly Group, October 2008

Figure 7

tion with 10GbE ports connected to Sun X2200MS servers and switch uplink ports to BlueArc's Titan. Before every test, all server/storage cache memories were cleared for the accurate and consistent test results. Engineers did not measure the performance number during ramp-up time but the number in the steady-state.

All tests were run twice and their results were averaged for the final result.

Terms of Usage

USE THIS DOCUMENT ONLY IF YOU AGREE TO THE TERMS LISTED HEREIN.

This document is provided, free-of-charge, to help you understand whether a given product, technology or service merits additional investigation for your particular needs. Any decision to purchase must be based on your own assessment of suitability.

This evaluation was focused on illustrating specific features and/or performance of the product(s) and was conducted under controlled, laboratory conditions and certain tests may have been tailored to reflect performance under ideal conditions; performance may vary under real-world conditions. Users should run tests based on their own real-world scenarios to validate performance for their own networks. Commercially reasonable efforts were made to ensure the accuracy of the data contained herein but errors and/or oversights can occur. In no event shall The Tolly Group be liable for damages of any kind including direct, indirect, special, incidental and consequential damages which may result from the use of information contained in this document.

The test/audit documented herein may also rely on various test tools the accuracy of which is beyond our control. Furthermore, the document relies on certain representations by the sponsor that are beyond our control to verify. Among these is that the software/hardware tested is production or production track and is, or will be, available in equivalent or better form to commercial customers.

When foreign translations exist, the English document is considered authoritative. To assure accuracy, only use documents downloaded directly from The Tolly Group's Web site.

All trademarks are the property of their respective owners.

208351-Cpvmfa2-cdb-03MAR09-verf