THE TOLLY GROUP

No. 208352 March 2009

BlueArc Corporation

Titan 3210 Network Storage System

Throughput Evaluation Using Common Internet File System (CIFS) 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 seamlessly support these 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 its Titan 3210 Network Storage System, a purpose-built network-attached storage (NAS) solution.

Tolly Group engineers measured the throughput of a single Titan 3210 node handling file-access requests using the Common Internet File System (CIFS) protocol. The back-end storage system used for testing was a Hitachi Data Systems (HDS) USP V outfitted with 256 Fibre Channel HDDs. (BlueArc and HDS are strategic partners delivering high-performance storage solutions.) Throughput was measured for a variety of read, write and mixed read/write operations from/to back-end disks with sequential access. Tolly also measured the throughput of two Titan 3210 nodes in a cluster in the same test scenarios.

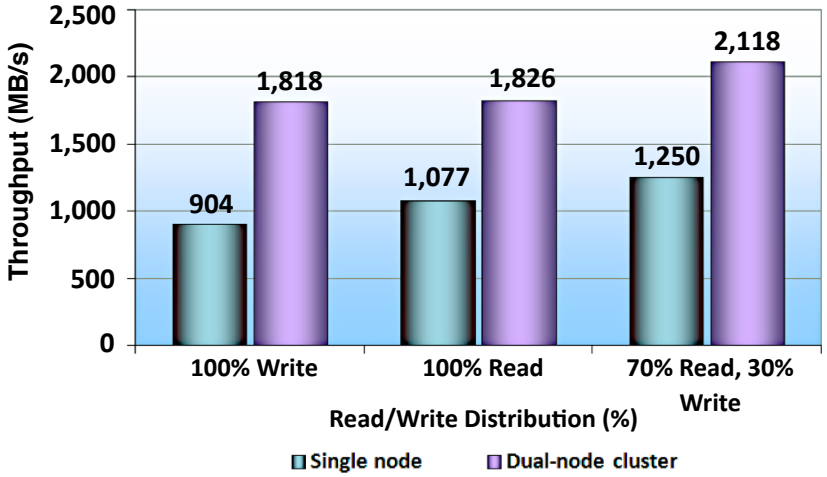
Tests were conducted in October 2008.

Test Highlights

- Delivers near wire-speed throughput when handling 100% sequential read operations over a 10GbE connection
- Provides linear scalability by achieving up to twice as much throughput with two nodes as compared to using a single node for all tests
- Sustains average throughput of up to 906 MB/s with a single-node and 1,812 MB/s for dual-node clusters in tests of 100% sequential writes for 32KB, 64KB and 128KB block sizes
- Proves that it can offer optimized, scalable high-performance NAS services for the CIFS environment running I/O intensive operations

Throughput for Various I/O Operations Using 64KB Block Size in Single-node and Dual-node Cluster Scenarios

as Reported by IOzone 3.283



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 shown here.

Source: The Tolly Group, October 2008

Figure 1

Executive Summary

The BlueArc Titan 3210 Network Storage System delivers near wire-speed, 10GbE throughput conducting sequential read operations for various block sizes using the CIFS file system and improves its performance by 1.6X to 2X in the same tests by clustering two nodes.

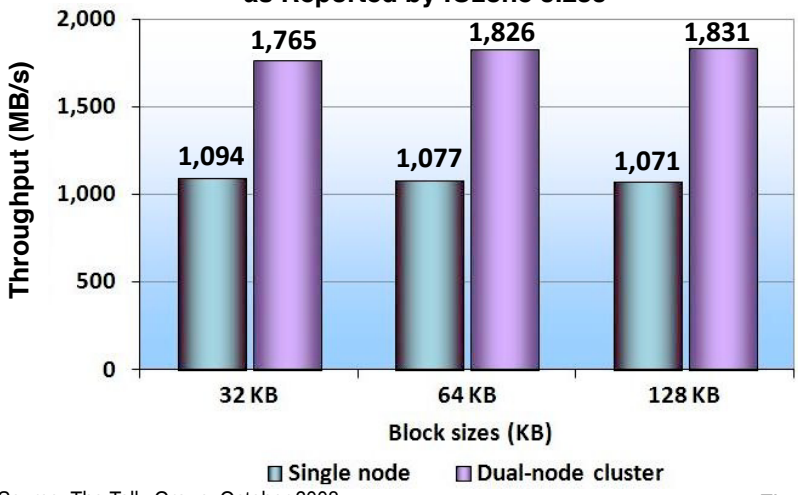
With the rapid increase in the amount of data handled by today's enterprise infrastructures, faster and more scalable storage systems are required to keep pace.

Conventional, general-purpose network-attached storage (NAS) systems depend upon CPUs to move data serially through the system and simultaneously manage cluster traffic. As additional filers are added to meet increased performance demands, the CPUs reach throughput limits and struggle with added cluster overhead. Overall efficiency is reduced as workload increases.

Conversely, the BlueArc hardware-based NAS architecture leverages Field Programmable Gate Arrays (FPGAs) to perform operations in parallel. This facilitates high throughput between servers and a dedicated storage backed with massive scalability.

This test focused on measuring throughput of a BlueArc Titan 3210 hardware-based NAS solution that accessed CIFS file system using different block sizes in a mix of read/write operations, and comparing that with throughput from a dual-node cluster. CIFS is one of the two most commonly used file systems in NAS solutions, along with NFS, and is the standard file sharing mechanism in Windows scenarios.

100% Sequential Read 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 2

The Tolly Group's hands-on evaluation of the BlueArc Titan 3210 demonstrated that the Titan 3210 delivered an average 1,080 MB/s (close to 10 Gbps considering network overhead) of throughput across 32KB, 64KB and 128KB block sizes when handling 100% sequential read operations using a single node and 1,810 MB/s (close to 16 Gbps) using a dual-node cluster. During 100% write operation, the Titan 3210 delivered about 910 MB/s and 1,810 MB/s with a single node and dual-node cluster, respectively. During a 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 proves that a single Titan 3210 offers excellent throughput using CIFS as a file-sharing mechanism in Windows environments. This also proves that its enhanced protocol acceleration helps to achieve the full potential of CIFS performance by optimizing its file processing. Not only does this mean that users can have better file sharing experiences but also the users can consolidate other NAS systems into a high-performance NAS, and support more enterprise application servers. When two Titan 3210 nodes are clustered, the tandem delivers even more impressive performance by supporting 1.6X to 2X the single-node performance without showing a significant sign of degradation, which is normally caused by the cluster overhead.

Results

100% Sequential Read Operation

Test results show that during 100% sequential read operation using a single node, the Titan 3210 achieved I/O throughput of 1,094 MB/s, 1,077 MB/s and 1,071 MB/s for 32KB, 64KB and 128KB block sizes, respectively. These convert to 8.8 Gbps, 8.6 Gbps and 8.6 Gbps of throughput. These throughput values represent near wire-speed performance over

a 10GbE connection. (Note: A default MTU of 1,500 bytes was used for the test.)

When clustering two nodes, engineers validated that throughput was improved dramatically by measuring 1.6X, 1.7X and 1.7X as much as the single node for 32KB, 64KB and 128KB block sizes. See Figure 2.

Regarding caching ...

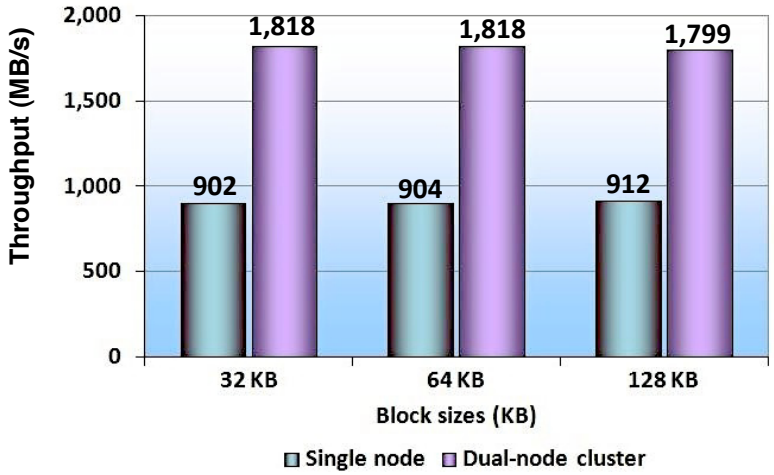
While caching data will help performance and is recommended in real-world deployments, our focus in this test was to push actual storage systems to their limits by forcing data to be retrieved from the actual disks. This test also took advantage of caching but, engineers deliberately configured cache 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 32GB per node even though the tested system was outfitted with 256GB cache memory. The caching configuration of the Titan 3210 remained in its default state.

This effort cannot be complete without the appropriate choice of a test data set. In this 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 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.

100% Sequential Write Operation

Test results show 902 MB/s, 904 MB/s, 912 MB/s throughput for 32KB, 64KB and 128KB block sizes, respectively. This represents a decrease of approximately 15% to 17% from the 100% read throughput.

100% Sequential 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 3*

BlueArc Corporation

Titan 3210 Network Storage System

Throughput Evaluation Using Common Internet File System (CIFS) 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
- Multi-node NAS cluster name space supporting up to 4 PB of unified directory space
- 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
Sales Inquiries
sales@bluearc.com
+1 866-864-1040

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

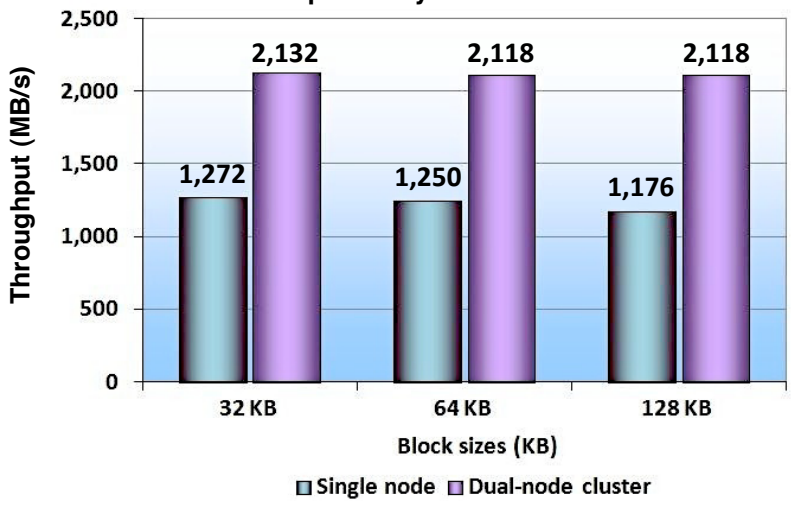
The Titan 3210's writing operation takes advantage of 4 GB of NVRAM inside one of the File System Modules (FSB). The NVRAM stores "writes" and returns fast acknowledgments to clients. This allows the clients to go on processing without waiting for the data to be written to disk. When NVRAM utilization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

With dual-node cluster configuration, the Titan 3210 reported throughput of 1,818 MB/s, 1,818 MB/s and 1,799 MB/s for 32, 64 and 128 KB block sizes, respectively, which is almost 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 test the system's mixed workload capability. As expected, the aggregate I/O throughput exceeded the theoretical maximum of one 10GbE connection by measuring 1,272 MB/s, 1,250 MB/s and 1,176 MB/s for 32KB, 64KB and 128KB block sizes, respectively, in a single-node scenario. This is almost equivalent to 11 Gbps of throughput on the Ethernet physical layer. The dual-node cluster results show throughput of 2,132 MB/s, 2,118 MB/s and 2,118 MB/s, which represents 1.7X, 1.7X and 1.8X the single-node performance for 32KB, 64KB and 128KB block sizes. See Figure 4.

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 the BlueArc Titan 3210 NAS system, client machines, a Force 10 Networks S2410P 10GbE LAN switch, a Brocade Silkworm 5320 SAN switch and HDS USP V storage system. See Figure 5 for the detailed product information and Figure 7 for the test bed diagram.

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

A total of 12 client machines (six per node) were used for the test. Each client ran the 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 in order to run IOzone iteratively with varying options.

For a single-node test, six client machines were used and IOzone was configured to use 32/64/128 KB 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 clients were used resulting in 96 threads. Everything else was the same as the single-node test.

The Brocade Silkworm 5320 SAN switch was configured to connect between Titan 3210's FC ports and HDS USP V's FC ports with a one-to-one connection. Each port-pair was grouped into a separate zone, yielding eight zones for the single-node test. There was only one zone configured for the dual-node cluster test so that all 32 FC ports were in the same zone.

Detailed Component Summary of BlueArc High-performance Storage Solution Tested

Vendor	Model	Description	Version	Quantity
BlueArc Corporation	**Titan 3210**	**Network Attached Storage (NAS) platform** with two 10GbE ports, eight 4Gbit/s FC 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 8 Back-end Directors (BEDs). Each FED has eight 4 Gbit/s FC ports. One FC 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:** Microsoft Windows 2003 Server R2 SP2 Enterprise Edition 64-bit for CIFS file service		12
Chelsio Communications	**S310E-SR+**	**10GbE Storage Accelerator** - a single 10GbE 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 4Gbit/s for this test)	6.1.0a	1
Force10 Networks	**S2410P**	**Data Center Ethernet Switch** supporting 24 10GbE ports	2.4.1.11	1

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

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**	902	1,818
	64 KB	904	1,818
	128 KB	912	1,799
100% Read	**32 KB**	1,094	1,765
	64 KB	1,077	1,826
	128 KB	1,071	1,831
70% Read / 30% Write	**32 KB**	1,272	2,132
	64 KB	1,250	2,118
	128 KB	1,176	2,118

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

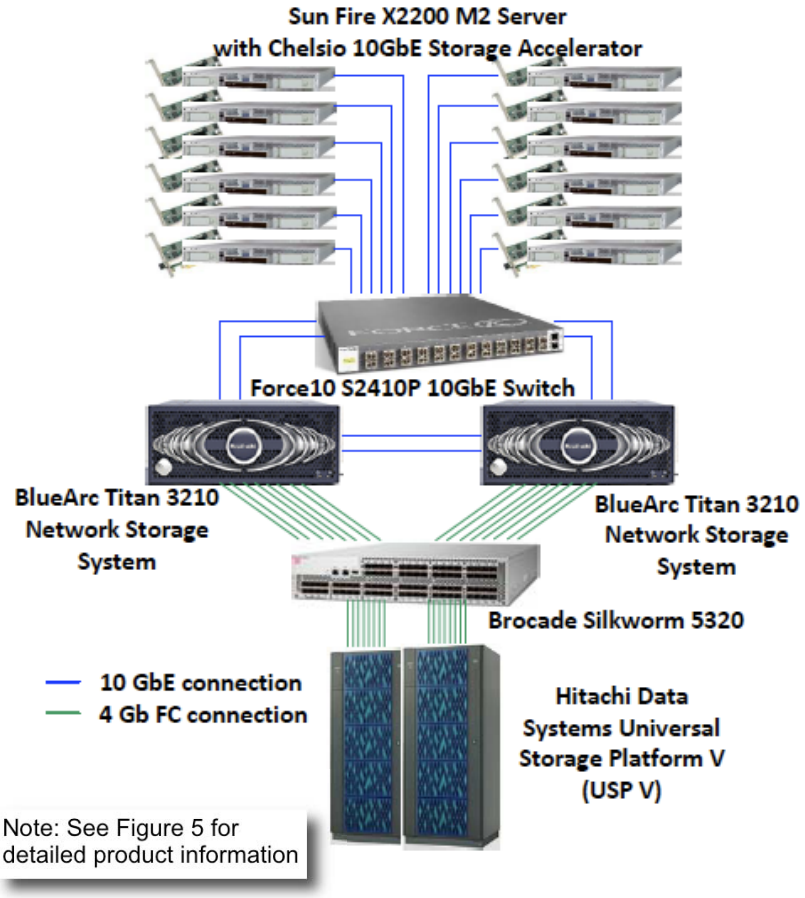
The HDS USP V was configured to use 32 GB of cache memory per node (64 GB for two nodes). Its 256 FC HDDs were configured for RAID 10 and 128 HDDs were supported for each node.

For network configuration, standard Ethernet MTU (1,500 bytes) was used. No Jumbo Frames were used for this test. The Force10 Networks S2410P 10GbE switch remained in its factory default configuration.

Before every test began, all server/storage cache memories were cleared to ensure accurate and consistent test results.

All tests were run twice and their results were averaged.

Test Bed Diagram
HDS High-performance Storage Solution



Source: The Tolly Group, October 2008 *Figure 7*

The Tolly Group is a leading global provider of third-party validation services for vendors of IT products, compo-nents and services.

The company is based in Boca Raton, FL and can be reached by phone at (561) 391-5610, or via the Internet at:

Web: http://www.tolly.com,
E-mail: sales@tolly.com

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 net-works. 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 conse-quential 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. 208352-Cpvmfa2-cdb-03MAR09-verf