
Arista 10 Gigabit Ethernet Switch Lab-Tested with Panasas ActiveStor Parallel Storage System Delivers Best Results for High-Performance and Low Latency for Scale-Out Cloud Storage Applications

Introduction

Ever-increasing growth in the volume of stored data coupled with applications that require demanding computational and processing power are driving industries such as finance, media, climate modeling, energy exploration, manufacturing design, computational fluid dynamics and physics, genetic research and Web services to deploy highly scalable cloud storage and cloud computing architectures to stay competitive.

10 Gigabit Ethernet prevails as the interconnect technology of choice for these applications. With its full non-blocking throughput, record density, low latency, ease of operation and leading TCO, the Arista 7100 Series 10 Gigabit Ethernet switch is ideal for storage applications. When combined with the Panasas ActiveStor Parallel Storage System, based on the ultra high-performance parallel, global file system architecture of the Panasas ActiveScale File System, the Arista 7100 Series 10Gb Ethernet switch provides a scalable, cost-effective cloud storage solution for a wide variety of high-performance and cloud computing applications.

This paper discusses the performance results obtained using the Panasas ActiveStor Parallel Storage System with the high-density 48-port Arista 7148SX fully non-blocking, low-latency L2/L3/L4 Ethernet switch.

Need for ever-increasing storage scalability

The increasing demand to support applications requiring high IOPs or large “bulk” storage of unstructured data is stretching the capability of existing architectures in the datacenter. In Oil & Gas, the ability to leverage HPC and storage for rapid analysis of large, complex seismic surveys influences capital decisions for oil and gas exploration in today's energy markets. Financial institutions rely upon complex time-series analysis that requires fast access to historical data and extreme low-latency when processing new market entries. Web services need the ability to scale rapidly to support ever-growing amounts of user-generated content each day. What all these companies have in common is a concrete need for ever-increasing scalability of their storage and processing capabilities in a cost-effective fashion.

Arista and Panasas provide a unique Cloud Storage solution

Arista's high-throughput, low latency 10Gb Ethernet switch coupled with the Panasas ActiveStor Parallel Storage System is a powerful combination of networking and storage that helps enterprises and grid computing environments eliminate bottlenecks and increase I/O throughput for their always-on and always growing environments. Unlike other high-performance storage systems that require the customer to build and maintain the storage system, Panasas has built a simple, easy to install, automatic to maintain storage system that delivers enterprise class data management features like snapshots and replication to lower the total cost of owning high performance storage. Panasas delivers high reliability as well as high performance so that mission critical applications like risk analysis for financial services firms or optical proximity correction for semiconductor companies. Together the combination of Arista and Panasas provide the performance, reliability and low cost of ownership that gives enterprises and cloud providers the best storage and network solution for scale-out applications.

Test Setup Description

The Panasas Lab used the Arista 7148SX, 48-port, 10 Gigabit Ethernet switch as a physical transport layer for both:

- The Panasas ActiveStor cluster network and
- The Panasas file-serving network used for client connectivity.

Testing was conducted utilizing the I/Ozone bandwidth benchmark (version 3.303) to profile the Arista/Panasas infrastructure for:

- File-serving bandwidth
- Single-stream performance
- Comparison to Infiniband infrastructure
- Network latency

The testing configuration consisted of the following¹:

Clients: (8) servers dual quad-core Intel Nehalem CPUs, 24GB memory, and Intel Oplink 10GbE adapter. Clients ran Linux 2.6.18-128 (CentOS 5.3) and accessed the Panasas ActiveScale File System through the DirectFLOW 3.4.0 protocol.

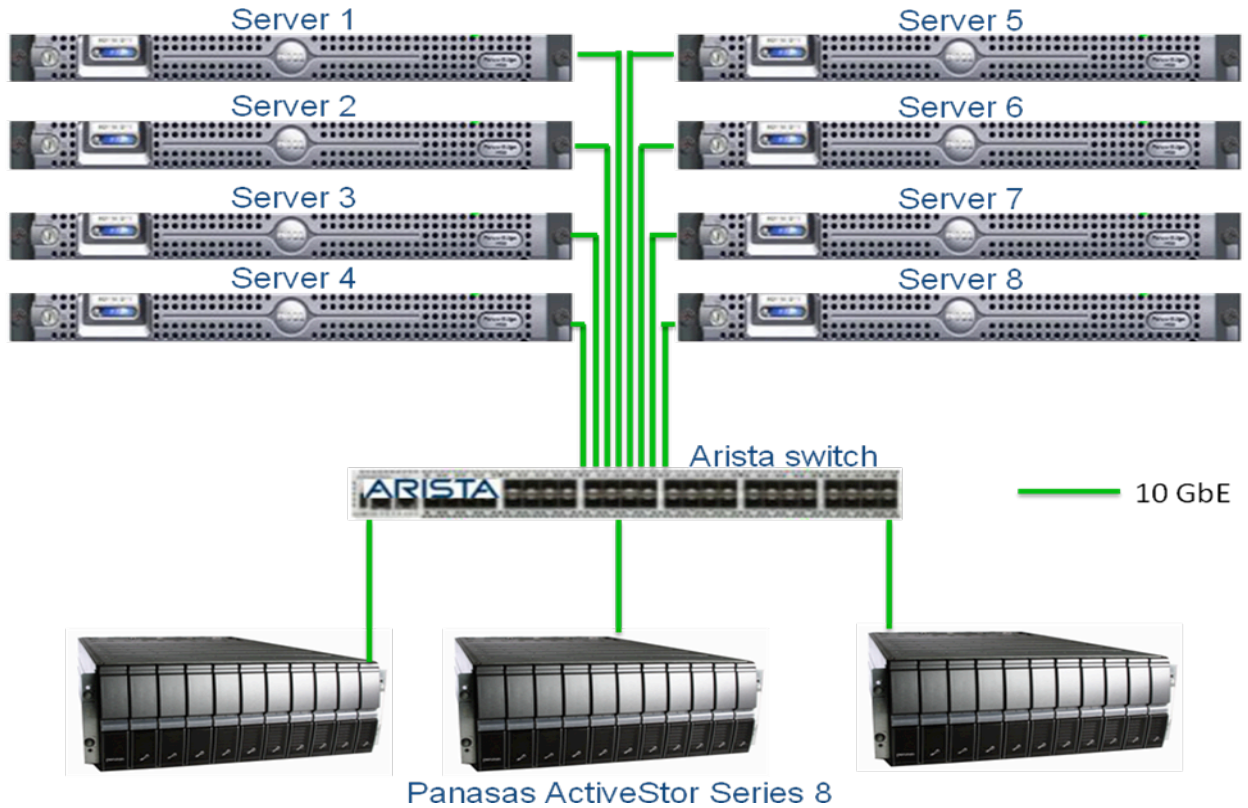
Panasas Storage: 3-shelf Series 8 Panasas storage cluster with 1+10 configuration using SB1000 blades for 30TB raw storage.

All file system clients were connected to the Arista switch via Intel Oplink 10GbE adapters.

The 10GbE infrastructure is utilized for both the client network NFS/CIFS traffic and the Panasas cluster network for communication between the Panasas servers.

The I/Ozone options included the sync() and close() time for the tests. Each client was writing/reading 48GB of data (2x physical memory) to prevent client side caching. Effects.

¹ The testing methodology describes the evaluation of the 10GbE technology and some concrete performance results. It serves as a reference point for choosing a scalable interconnect technology for high performance and high bandwidth applications. It is not meant to be used as a definitive guide to 10GbE and Panasas.



Throughput Testing

Testing the throughput consisted of measuring the rate of storage-to-client reads and client-to-storage writes as follows:

1. The client (Client 1) reads from the Panasas PAS8 storage.
2. The client (Client 1) writes to the Panasas PAS8 storage (See Figure below for results.)

The results with 10 GbE are shown in the table below:

Test: Clients	Test: Threads/Client	Test: Shelves	Reads (MB/sec)	Writes (MB/sec)
1	1	3	442	356
1	8	3	955	841
8	3	1	710	669
8	6	2	1221	1213
8	8	3	1869	1779

We can determine several key points from these tests:

1. For an application performing single-threaded or serial I/O, we saw strong throughput at 356 MB/S for writes, and over 400 MB/S for reads.
2. Applications that can take advantage of multiple I/O threads on a single client can see significantly more performance at up to 841 MB/S for writes, and 955 MB/S for reads. With a maximum theoretical limit of 1000 MB/S for the single 10 GbE link to the client, this is excellent.
3. When increasing the number of clients and the number of shelves, the I/O performance scales in an optimal linear fashion, rather than hitting a plateau or declining with increasing load.

Comparing to InfiniBand Throughput

Throughput was also measured using an Infiniband infrastructure, in lieu of the Arista 10 Gigabit Ethernet switch and Intel Oplink 10 Gigabit Ethernet adapters. The InfiniBand infrastructure is tested at DDR rates - although the results can be expected to be virtually identical at QDR rates, as the network is not the bottleneck for either 10 GbE or InfiniBand.

The table below compares throughput speed for typical file system operations for a single client using a single thread to access the Panasas ActiveStor Parallel Storage System running over 1 Gigabit Ethernet, 10 Gigabit Ethernet and Infiniband.

Operation	1GbE (MB/sec)	10GbE (MB/sec)	InfiniBand (MB/sec)
Panasas read:	96	442	465*
Panasas write:	83	356	335

*Showing beneficial read caching effects due to insufficient file size. Without this effect, actual direct comparison results would be somewhat lower for the InfiniBand read operation.

We can learn several important things from these data:

1. In the high-performance Panasas parallel storage environment, the Arista 10 GbE solution and an InfiniBand solution both deliver close to 5 times the performance of 1 GbE interconnects for reads.
2. These same solutions deliver over 4 times the throughput versus 1 GbE for writes.
3. Performance of the 10 GbE solution versus InfiniBand shows a less than 5% advantage to IB for reads, and a slightly more than 5% advantage to 10 GbE for writes.

Ethernet environments have the distinct advantage of tools like ping and traceroute, which contribute to making these environments easier to manage for network operations. This bypasses the need to re-train staff, and ensures a faster deployment cycle for 10 GbE in HPC environments for network staff trained in IP and Ethernet technology.

Summary

These performance results, generated by combining Arista's 7148SX and the Panasas ActiveStor Parallel Storage System, clearly show the significant benefits of using 10 Gigabit Ethernet versus 1 Gigabit Ethernet or Infiniband interconnect. For today's high performance computing environments, the Arista and Panasas solution at 10 GbE delivers better cost/performance than 1 GbE, while providing significantly greater throughput. As well, this solution compares favorably to InfiniBand, as performance levels are very similar, yet the 10 GbE solution leverages customer investments in standard Ethernet tools and technology, without requiring application rewrites and the requirement to support an additional niche technology in the data center. In particular, coupling the Panasas ActiveStor Parallel Storage System with Arista high-density, low-latency 10 GbE switches yields near optimum results on 10Gb Ethernet for applications that can take advantage of multiple I/O threads on a single client.

This unique combination of networking and storage provides the ability to smoothly scale storage capacity and bandwidth requirements for customers in many industries, including Oil & Gas, Financial Services, Media and Entertainment, and Web Services.

About Arista

The Arista 7100 Series of datacenter Ethernet switches feature the industry's highest density 10 Gigabit Ethernet switching solution and the first with an extensible modular network operating system. With breakthrough price-performance, the Arista 7100 Series enables 10 Gigabit Ethernet to be deployed throughout in the data center, which can significantly improve server utilization and data center power efficiency.

At the core of Arista's platform is the Extensible Operating System (EOS™), pioneering new software architecture with self-healing and live in-service software upgrade capabilities. EOS leapfrogs existing network OS designs, by providing the following capabilities and benefits:

- In-service software updates – enable reduced maintenance windows due to ability to update processes without system interruption.
- Software fault containment – all faults are contained in a single module, thus providing superior system stability.
- Stateful fault repair – continuous health monitoring of all processes enables invisible repair of faults.
- Security exploit containment – EOS modular architecture improves security by limiting any potential vulnerability to an individual module.
- Scalable management interface – enables automated maintenance, updates, and integration with third party network management systems.

Arista's high throughput, low latency switch, with the proprietary EOS architecture, is the most advanced and reliable 10Gb Ethernet switch in the industry today.

For more information, visit <http://www.aristanetworks.com>.

About Panasas

Panasas, Inc., the leading provider of high-performance storage for the world's most performance-intensive applications, helps commercial, government and academic organizations accelerate their time to results leading to real world breakthroughs that improve people's lives. Panasas' high-performance storage systems enable customers to maximize the benefits of compute clusters by eliminating the storage bottleneck created with legacy network storage technologies. Panasas ActiveStor Parallel Storage Clusters, in conjunction with the ActiveScale® Operating Environment and PanFS™ parallel file system, offers the most comprehensive portfolio of storage solutions for high performance computing (HPC) environments. Panasas is headquartered in Fremont, California. **For more information, visit www.panasas.com.**

Panasas, ActiveScale, DirectFLOW and PanFS are trademarks or registered trademarks of Panasas, Inc. All other trademarks are the property of their respective owners.