

Arista 10 Gigabit Ethernet Switch Lab Tested with IBRIX Fusion™ Software Deliver 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, 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, and leading TCO, the Arista 7100 Series 10 Gigabit Ethernet switch is ideal for storage applications. When combined with IBRIX Fusion™, a purpose-built scale-out NAS solution, 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 IBRIX Fusion scale-out NAS solution with the Arista 7124S fully non-blocking, low-latency L2/L3 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 media and entertainment, animation rendering and special effects requires processing tens of thousands of files to compose one second of film. 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 order to keep their businesses running without service loss.

Arista and IBRIX provide a unique Cloud Storage solution

Arista’s high-throughput, low latency 10Gb Ethernet switch coupled with the IBRIX scalable file serving NAS solution 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. IBRIX Fusion meets the challenges of today’s high performance storage architectures with its patented segmented file system™ which improves data delivery from the storage environment to the computing elements. The software-based solution can be deployed with any market-leading servers, storage and network interconnect technologies like Arista’s 10Gb Ethernet switches, providing customers a flexible deployment methodology for the best performance at the right price. IBRIX allows increased storage efficiency through flexible hardware topologies and advanced software features, including data tiering, remote replication and dynamic load balancing.

Test Setup Description

The IBRIX Lab used the Arista 7124S, 24- port, 10 Gigabit Ethernet switch as a physical transport layer for both:

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

Testing was conducted over 5 days to profile the Arista/IBRIX infrastructure for:

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

The testing configuration consisted of the following¹:

Clients: (2) Sun x4150 servers – 2 quad core CPUs, 16GB memory, Chelsio 320 10GbE adapter. Clients ran RHEL5u2 Linux and accessed the IBRIX file system through NFS and IBRIX protocols.

File serving nodes (segment servers): (2) Dell 1950 PowerEdge™ servers – 2 quad core CPUs, 16GB memory, Chelsio 320 10GbE adapter, bonded 1GbE IP SAN ports.

Storage: (2) Dell EqualLogic PS5000X iSCSI arrays (6.4TB 400GB SAS Drives)

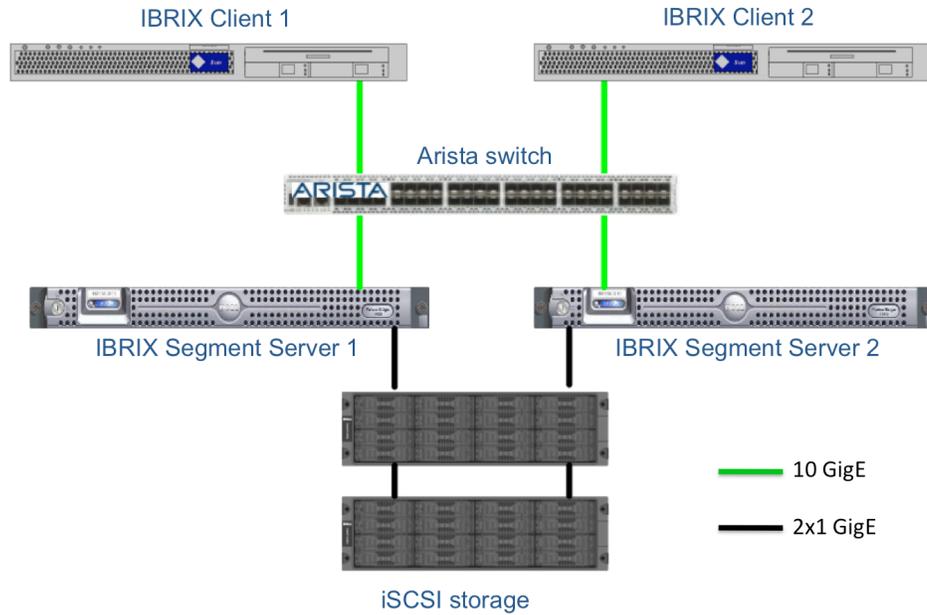
All file system clients and segment servers were connected to the Arista switch via N320 series Chelsio 10GbE adapters.

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

After installing the drivers and configuring the interfaces a streamlined process that takes less than 5 minutes per server and that is part of the standard IBRIX installation bundle, the hosts are able to communicate over the 10GbE network. (See Figure below.)

The segment servers utilize bonded 1GbE ports for connectivity to the IP SAN for the back-end iSCSI storage arrays.

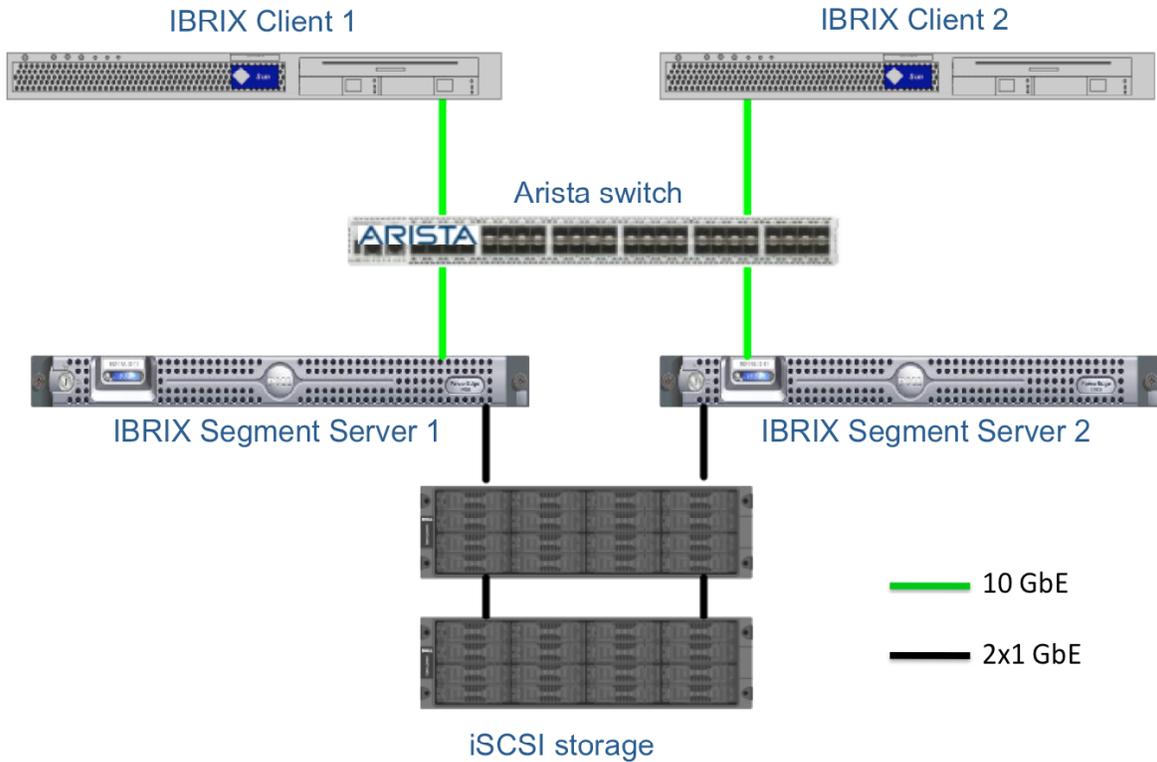
¹ The testing methodology describes the evaluation of the 10GbE technology and some concrete performance results. It serves as a reference point for choosing a scalable interconnect technology for high performance and high bandwidth applications and is not meant to be used as a definitive guide to 10GbE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



Throughput Testing

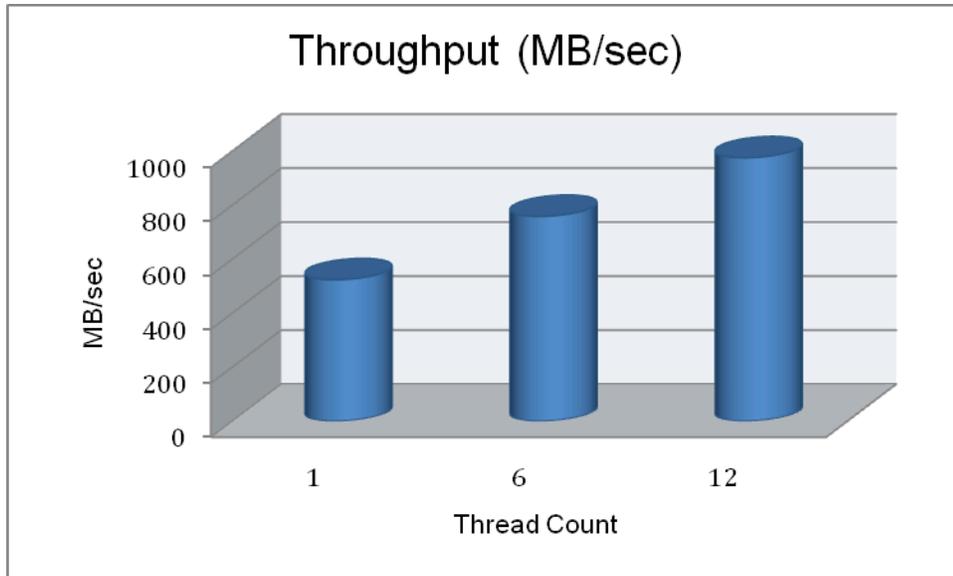
Testing the throughput consisted of measuring the rate of host-to-host reads as follows:

1. The first IBRIX client (Client 1) writes to the 1st IBRIX segment server (IBRIX Segment Server 1).
2. These files are cached in memory
3. The second IBRIX client (Client 2) reads these files. (See Figure below.)



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

Test	Results (MB/sec)	Notes
Single file read	524	This result is the largest single-threaded throughput that was ever achieved with IBRIX across any interconnect
6-thread test	759	These results are the highest achieved with IBRIX when utilizing a one-to-one client-to-segment server ratio
12-thread test	976	



Comparing to Infiniband Throughput

Throughput was also measured using an Infiniband infrastructure, in lieu of the Arista 10 Gigabit Ethernet switch and Chelsio 10 Gigabit Ethernet adapters.

The table below compares throughput speed for typical file system operations using IBRIX Fusion running over 10 Gigabit Ethernet and Infiniband.

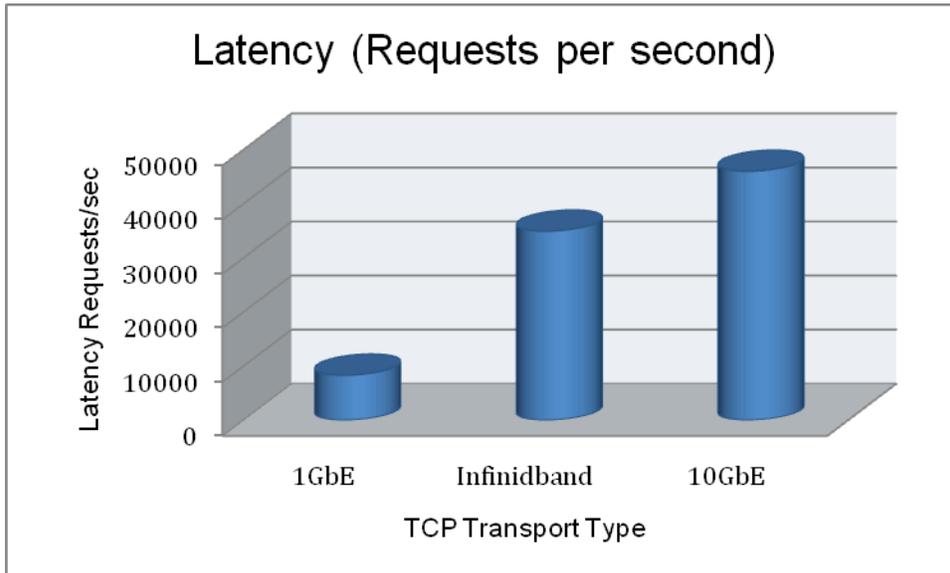
Operation	10GbE (MB/sec)	InfiniBand (MB/sec)
IBRIX read:	675	649
NFS read:	649	651
IBRIX read cached:	976	876
NFS read cached:	910	691
IBRIX write:	368	354
NFS write:	402	361

Latency Testing

For these tests, IBRIX used the netperf utility test, and obtained the following results:

- The TCP request/response test between the two hosts produces a throughput of 9.8 Gb/s, just shy of the maximum theoretical 10Gb/s capacity on the link.

- The request/response test across the Arista 7124S 10Gigabit Ethernet switch produces 46,082 requests and responses per second. This is in contrast to 8352.91 requests and responses per second over a 1 Gigabit Ethernet link. In other words, the request/response throughput over 10Gigabit Ethernet is 5.5 times higher than over standard 1Gigabit Ethernet (see Figure below).



Summary

These performance results, combining Arista's 7100 and IBRIX Fusion clearly shows the significant latency and performance benefits of 10 Gigabit Ethernet versus 1 Gigabit Ethernet or Infinidband interconnect. In particular, coupling IBRIX Fusion with Arista yields one of the best single-threaded results on 10Gb Ethernet.

This unique combination of networking and storage provides the ability to smoothly scale storage capacity and bandwidth requirements without loss of service for customers in many industries, including 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 IBRIX

IBRIX® delivers extreme data storage scalability and performance that helps enterprises cost effectively harness the unstructured data explosion. Internet service providers, content delivery networks, animation rendering and visual effects companies, banks and hedge funds rely upon IBRIX to increase application performance, enable storage scalability, simplify management, and deliver significant storage cost savings.

IBRIX's flagship product, IBRIX Fusion, is a scale-out NAS solution that differs from traditional NAS by allowing data access to be delivered through as many file servers as needed to achieve the required aggregate performance and capacity scaling. IBRIX's strength lies in its ability to deliver high volume, petabyte-level storage as well as high I/O performance. Unlike alternative hierarchical-based solutions which are prone to hot spots, limit scalability, and require more data hops to get to files, IBRIX Fusion has no central metadata component or distributed lock manager, allowing all data and metadata to be processed in parallel. The result is that capacity and performance can be independently scaled either by adding more disks to support high volume, unstructured data growth, or more servers for fast access to data-intensive applications. IBRIX Fusion allows the user to control costs by supporting its installation on any industry-standard servers and any combination of direct-attached storage (DAS), network-attached storage, or in a storage area network (SAN). In application clusters, which have large numbers of client nodes that perform millions of operations per second on the same files simultaneously, IBRIX Fusion relieves the file I/O bottlenecks inherent in this architecture by supporting the addition of multiple file servers, which are built using industry standard servers running Linux and IBRIX Fusion software. These file servers have the ability to cache files and deliver them in parallel to their respective clients. Additional file servers can be added as needed while the existing file system is in production and online.

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