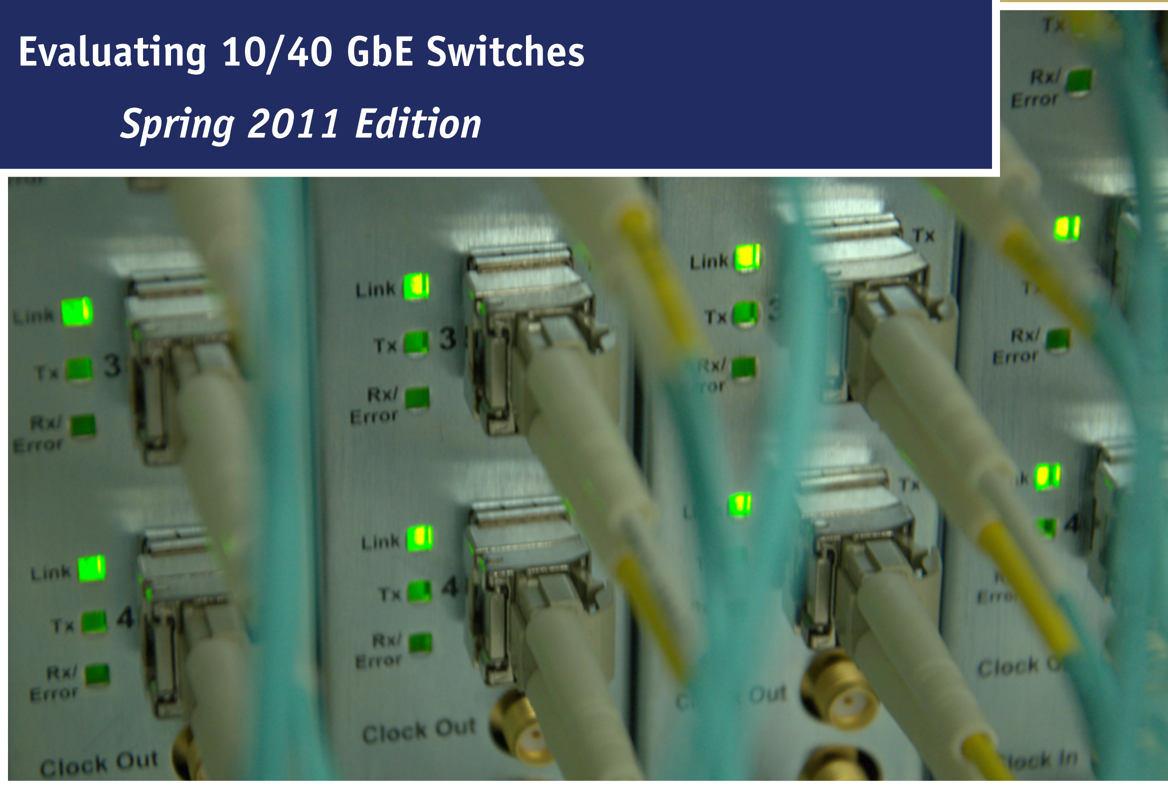
Lippis Report

Open Industry Network Performance & Power Test
for
Cloud Networks

Evaluating 10/40 GbE Switches

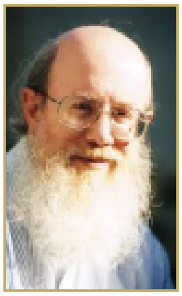
Spring 2011 Edition



May, 2011

© Lippis Enterprises, Inc. 2011

Forward by Scott Bradner



It is now almost twenty years since the Internet Engineering Task Force (IETF) initially chartered the Benchmarking Methodology Working Group (bmwg) with me as chair. The aim of the working group was to develop standardized terminology and testing methodology for various performance tests on network devices such as routers and switches.

At the time that the bmwg was formed, it was almost impossible to compare products from different vendors without doing testing yourself because each vendor did its own testing and, too often, designed the tests to paint their products in the best light. The RFCs produced by the bmwg provided a set of standards that network equipment vendors and testing equipment vendors could use so that tests by different vendors or different test labs could be compared.

Since its creation, the bmwg has produced 23 IETF RFCs that define performance testing terminology or methodology for specific protocols or situations and are currently working on a half dozen more. The bmwg has also had three different chairs since I resigned in 1993 to join the IETF's steering group.

The performance tests in this report are the latest in a long series of similar tests produced by a number of testing labs. The testing methodology follows the same standards as I was using in my own test lab at Harvard in the early 1990s thus the results are comparable. The comparisons would not be all that useful since I was dealing with far slower speed networks, but a latency measurement I did in 1993 used the same standard methodology as do the latency tests in this report.

Considering the limits on what I was able to test way back then, the Ixia test setup used in these tests is very impressive indeed. It almost makes me want to get into the testing business again.

Scott is the University Technology Security Officer at Harvard University. He writes a weekly column for Network World, and serves as the Secretary to the Board of Trustees of the Internet Society (ISOC). In addition, he is a trustee of the American Registry of Internet Numbers (ARIN) and author of many RFC network performance standards used in this industry evaluation report.

License Rights

© 2011 Lippis Enterprises, Inc. All rights reserved. This report is being sold solely to the purchaser set forth below. The report, including the written text, graphics, data, images, illustrations, marks, logos, sound or video clips, photographs and/or other works (singly or collectively, the "Content"), may be used only by such purchaser for informational purposes and such purchaser may not (and may not authorize a third party to) copy, transmit, reproduce, cite, publicly display, host, post, perform, distribute, alter, transmit or create derivative works of any Content or any portion of or excerpts from the Content in any fashion (either externally or internally to other individuals within a corporate structure) unless specifically authorized in writing by Lippis Enterprises. The purchaser agrees to maintain all copyright, trademark and other notices on the Content. The Report and all of the Content are protected by U.S. and/or international copyright laws and conventions, and belong to Lippis Enterprises, its licensors or third parties. No right, title or interest in any Content is transferred to the purchaser.

Purchaser Name:

Acknowledgements

We thank the following people for their help and support in making possible this first industry evaluation of 10GbE private/public data center cloud Ethernet switch fabrics.

Atul Bhatnagar, CEO of Ixia, for that first discussion of this project back at Interop in May 2009 and his continuing support throughout its execution.

Scott O. Bradner of Harvard University for his work at the IEFT on network performance benchmarks, used in this project, and the Forward of this report.

Leviton for the use of fiber optic cables equipped with optical SFP+ connectors to link Ixia test equipment to 10GbE switches under test.

Michael Githens, Lab Program Manager at Ixia, for his technical competence, extra effort and dedication to fairness as he executed test week and worked with participating vendors to answer their many questions.

Henry He, Technical Product Manager, Kelly Maloit, Director of Public Relations, Jim Smith, VP of Marketing, all from Ixia, for their support and contributions to creating a successful industry event.

Bob Anderson, Photography and Videography at Ixia, for video podcast editing.

All participating vendors and their technical plus marketing teams for their support of not only the event but to each other, as many provided helping hands on many levels to competitors.

Tiffany O'Loughlin for her amazing eye in capturing test week with still and video photography.

Bill Nicholson for his graphic artist skills that make this report look amazing.

Jeannette Tibbetts for her editing that makes this report read as smooth as a test report can.

Steven Cagnetta for his legal review of all agreements.

© Lippis Enterprises, Inc. 2011

Table of Contents

© Lippis Enterprises, Inc. 2011

Executive Summary

To assist IT business leaders with the design and procurement of their private or public data center cloud fabric, the Lippis Report and Ixia have conducted an open industry evaluation of 10GbE data center switches. In this report, IT architects are provided the first comparative 10 Gigabit Ethernet Switch (10GbE) performance information to assist them in purchase decisions and product differentiation.

The resources available for this test at Ixia's iSimCity are out of reach for nearly all corporate IT departments with test equipment on the order of $9.5M, devices under test on the order of $2M, plus costs associated with housing, power and cooling the lab, and lastly nearly 22 engineers from around the industry working to deliver test results. It's our hope that this report will remove performance, power consumption and latency concern from the purchase decision, allowing IT architects and IT business leaders to focus on other vendor selection criteria, such as post sales support, platform investment, vision, company financials, etc.

The Lippis test reports based on independent validation at Ixia's iSim City, communicates credibility, competence, openness and trust to potential buyers of 10GbE data center switching equipment as the tests are open to all suppliers and are fair, thanks to RFC and custom-based tests that are repeatable. The private/public data center cloud 10GbE fabric test was free for vendors to participate and open to all industry suppliers of 10GbE switching equipment, both modular and fixed configurations.

The tests took place during the week of December 6 to 10, 2010 and April 11 to 15, 2011 in the modern Ixia test lab, iSimCity, located in Santa Clara, CA. Ixia supplied all test equipment needed to conduct the tests while Leviton provided optical SPF+ connectors and optical cabling. Each 10GbE supplier was allocated lab time to run the test with the assistance of an Ixia engineer. Each switch vendor configured their equipment while Ixia engineers ran the test and logged the resulting data.

The tests conducted were IETF RFC-based performance and latency test, power consumption measurements and a custom cloud-computing simulation of large north-south plus east-west traffic flows. Both Top-of-Rack (ToR) and Core switches were evaluated.

ToR switches evaluated were:

Arista 7124SX 10G SFP Data Center Switch

Arista 7050S-64 10/40G Data Center Switch

BLADE Network Technologies, an IBM Company
IBM BNT RackSwitch G8124

BLADE Network Technologies, an IBM Company
IBM BNT RackSwitch G8264

Brocade VDX™ 6720-24 Data Center Switch

Force10 S-Series S4810

Hitachi Cable, Apresia15000-64XL-PSR

Voltaire® Vantage™ 6048

Core switches evaluated were:

Alcatel-Lucent OmniSwitch 10K

Arista 7504 Series Data Center Switch

Juniper Network EX Series EX8200 Ethernet Switch

From the above list the following products were tested in April, 2011 and added to the products evaluated in December 2010.

- Arista 7124SX 10G SFP Data Center Switch
- Arista 7050S-64 10/40G Data Center Switch
- Brocade VDX™ 6720-24 Data Center Switch

© Lippis Enterprises, Inc. 2011

The following list our top ten findings:

1. 10GbE ToR and Core switches are ready for mass deployment in private/public data center cloud computing facilities, as the technology is mature, and products are both stable and deliver stated goals of high performance, low latency and low power consumption.

2. There are differences between suppliers, and it's recommended to review each supplier's results along with other important information in making purchase decisions.

3. All ToR and Core switches evaluated in the Lippis/Ixia test were recently introduced to market where the Lippis/Ixia test was their first open public evaluation.

4. Most ToR switches are based upon a new generation of merchant silicon that provides a single chip-forwarding engine of n-10GbE ports that delivers consistent performance and low power consumption.

5. All ToR and Core switches offer low power consumption with energy cost over a three-year period estimated between 1.3% and 4% of acquisition cost. We measure WATTS per 10GbE port via ATIS methods and TEER values for all switches evaluated. It's clear that these products represent the state-of-the-art in terms of energy conservation. ToR switches consume between 3.58 and 5.5 WATTS per 10GbE port while Core switches consume between 10.3 and 21.68 WATTS per 10GbE port.

6. Most switches design airflow from front-to-back or back-to-front rather than side-to-side to align with data center hot/cold aisle cooling design.

7. There are differences between suppliers in terms of the network services they offer, such as virtualization support, quality of service, etc., as well as how their Core/Spine switches connect to ToR/Leaf switches to create a data center fabric. It's recommended that IT business leaders evaluate Core switches with ToR switches and vice versa to assure that the network services and fabric attributes sought after are realized throughout the data center.

8. Most Core and ToR switches demonstrated the performance and latency required to support storage enablement or converged I/O. In fact, all suppliers have invested in storage enablement, such as support for Converged Enhanced Ethernet (CEE), Fibre Channel over Ethernet (FCoE), Internet Small Computer System Interface (iSCSI), Network-attached Storage (NAS), etc., and while these features were not tested in the Lippis/Ixia evaluation, most of these switches demonstrated that the raw capacity is built into the switches for its support.

9. Two ToR switches support 40GbE while all Core switches possess the backplane speed capacity for some combination of 40 and 100GbE. Core switches tested here will easily scale to support 40GbE and 100GbE while ToR suppliers offer uplinks at these speeds.

10. The Lippis/Ixia test demonstrated the performance and power consumption advantages of 10GbE networking, which can be put to work and exploited for corporate advantage. For new server deployments in private/public data center cloud networks, 10GbE is recommended as the primary network connectivity service as a network fabric exists to take full advantage of server I/O at 10GbE bandwidth and latency levels.

© Lippis Enterprises, Inc. 2011

Market Background

The following market background section provides perspective and context to the major changes occurring in the IT industry, and its fundamental importance to the network switches that were evaluated in the Lippis/Ixia test.

The IT industry is in the midst of a fundamental change toward centralization of application delivery via concentrated and dense private and public data center cloud computing sites. Corporations are concentrating IT spending in data centers and scaling them up via server virtualization to reduce complexity and associated operational cost while enjoying capital cost savings advantage. In fact, there is a new cloud-o-nomics that is driven by the cost advantage of multi-core computing coupled with virtualization that has enabled a scale of computing density not seen before. Couple this with a new tier of computing, Apple's iPad, Android tablets plus smartphones that rely upon applications served from private and public cloud computing facilities, and you have the making of a systemic and fundamental change in data center design plus IT delivery.

The world has witnessed the power of this technology and how it has changed the human condition and literally the world. On May 6, 2010, the U.S. stock market experienced a flash crash that in 13 short seconds, 27,000 contracts traded consisting of 49% of trading volume. Further, in 15 minutes $1 trillion of market value disappeared. Clusters of computers connected at 10GbE programmed to perform high frequency trades based upon proprietary algorithms did all this. The making of the movie *Avatar* was produced at WETA Digital with 90% of the movie developed via giant computer clusters connected at 10GbE to process and render animation. As big science cost has soared, scientists have moved to high performance computing clusters connected at 10GbE to simulate rather than build scientific experiments. Biotech engineers are analyzing protein-folding simulations while military scientists simulate nuclear explosions rather than detonate these massive bombs. Netflix has seen its stock price increase as it distributes movies and TV programs over the Internet and now threatens Comcast and others on-demand business thanks to its clusters of computers connected at 10GbE. At the same time, Blockbuster Video filed for bankruptcy.

A shift in IT purchasing is occurring too. During the market crash of 2008, demand for information increased exponentially while IT business decision makers were forced to scale back budgets. These diverging trends of information demand and IT budget contraction have created a data center execution gap where IT leaders are now forced to upgrade their infrastructure to meet information and application demand. To close the gap, IT business decision makers are investing in 10GbE switching. Over the past several quarters, sales of 10GbE fixed and modular switches have grown in the 60% to 70% range.

At the center of this new age in private and public data center cloud computing is networking, in particular Ethernet networking that links servers to storage and the internet providing high-speed transport for application traffic flow. As the data center network connects all devices, it is the single most important IT asset to assure end users receive an excellent experience and service is continually available. But in the virtualization/cloud era of computing, networking too is fundamentally changing as demand for low latency, high performance and lower power consumption switches increase to address radically different traffic patterns. In addition, added network services, such as storage enablement, are forcing Ethernet to become not just a connectivity service in the data center but a fabric that connects computing and storage; transporting data packets plus blocks of storage traffic. In short, IT leaders are looking for Ethernet to provide a single data center fabric that connects all IT assets with the promise of less equipment to install, simpler management and speed.

The following characterize today's modern private and public data center cloud network fabric.

Virtualization: The problems with networking in virtualized infrastructure are well documented. Move a virtual machine (VM) from one physical server to another

© Lippis Enterprises, Inc. 2011

and network port profile, Virtual Local Area Networking (VLAN), security settings, etc., have to be reconfigured, adding cost, delay and rigidity to what should be an adaptive infrastructure. Many networking companies are addressing this issue by making their switches virtualization aware, auto reconfiguring and/or recommending a large flat layer 2 domain. In addition, networking companies are offering network management views that cross administrative domains of networking and virtualization in an effort to provide increased visibility and management of physical and virtual infrastructure. The fabric also needs to eliminate the boundary between physical and virtual infrastructure, both in terms of management and visibility plus automated network changes as VMs move.

Traffic Patterns Shift with Exponential Volume Increase: Data centers are experiencing a new type of traffic pattern that is fundamentally changing network design plus latency and performance attributes. Not only are traditional north-south or client-server flows growing but east-west or server-server and storage-server flows now dominate most private and public data center cloud facilities. A dominate driver of east-west flows is the fact that the days of static web sites are over while dynamic sites take over where one page request can spawn 50 to 100 connections across multiple servers.

Large contributors to east-west flows are mobile devices. Mobile application use is expanding exponentially, thanks to the popularity of the iPhone, iPad and increasingly Android smartphones plus tablets. As of this writing, there are some 205,000 plus smartphone applications. The traffic load these applications are placing on data center Ethernet fabrics is paradoxically immense. The vast majority of mobile applications are hosted in data centers and/or public cloud facilities. The application model of mobile devices is not to load them with thick applications like Microsoft Word, PowerPoint, Excel, etc, but to load them with thin clients that access applications and data in private and/or public data centers, cloud facilities.

A relatively small query to a Facebook page, *New York Times* article, Amazon purchase, etc., sends a flurry of east-west traffic to respond to the query. Today's dynamic web spreads a web page over multiple servers and when called upon, drives a tsunami of traffic, which must be delivered well before the person requesting the data via mobile device or desktop loses interest and terminates the request.

East-west traffic will only increase as this new tier of tablet computing is expected to hit 19.5 million units in 2010 and 54.8 million in 2011, according to *The Wall Street Journal*. Add tens of millions of smartphones to the mix, all relying on private/public data center cloud facilities for their applications, and you have the ingredients for a sustained trend placing increased pressure for east-west traffic flows upon Ethernet data center fabrics.

Low Latency: Data experiences approximately 10 ns of latency as it traverses across a computer bus passing between memory and CPU. Could networking become so fast that it will emulate a computer bus so that a data center operates like one giant computer? The answer is no, but the industry is getting closer. Today's 10GbE switches produce 400 to 700 ns of latency. By 2014, it's anticipated that 100GbE switching will reduce latency to nearly 100 ns.

Special applications, especially in high frequency trading and other financial markets, measure latency in terms of millions of dollars in revenue lost or gained, placing enormous pressure on networking gear to be as fast as engineers can design. As such, low latency is an extremely important consideration in the financial services industry. For example, aggregating all 300 plus stock market feeds would generate approximately 3.6Gbs of traffic, but IT architects in this market choose 10GbE server connections versus 1GbE for the sole purpose of lower latency even though the bandwidth is nearly three times greater than what theoretically would be required.

More generally, latency is fundamental to assuring a user's excellent experience. With east-west traffic flows making up as much as 80% of data center traffic, low latency require-

© Lippis Enterprises, Inc. 2011

ments are paramount as every trip a packet makes between servers during the processes of responding to an end user's query adds delay to the response and reduces the user experience or revenue potential.

Storage Traffic over Ethernet Fabric: Converged I/O or unified networking where storage and network traffic flow over a single Ethernet network will increasingly be adopted in 2011. Multiple storage enablement options exist, such as iSCSI over Ethernet, ATA over Ethernet (AoE), FCoE, etc. For IP-based storage approaches, such as iSCSI over Ethernet and AoE, all that is needed is a 10GbE network interface card (NIC) in a server to support both storage and data traffic, but the Ethernet fabric requires lossless Ethernet to assure integrity of storage transport over the fabric as blocks of data transmit between server and storage farms.

For FCoE, a single converged network adaptor or CNA plugged into a server provides the conduit for storage and application traffic flows to traverse over an Ethernet fabric. The number of suppliers offering CNAs has grown significantly, including Intel, HP, Emulex, IBM, ServerEngines, QLogic, Cisco, Brocade, etc. In addition, the IEEE opened up the door for mass deployment as it has ratified the key Ethernet standards for lossless Ethernet. What will drive converged I/O is the reduced cost of cabling, NIC and switching hardware.

Low Power Consumption: With energy costs increasing plus green corporate initiatives as well as government mandates pervasive, all in an effort to reduce carbon emissions, IT leaders have been demanding lower energy consumption of all IT devices, including data center network equipment. The industry has responded with significant improvements in energy efficiency and overall lower wattage consumption per port while 10GbE switches are operating at full line rate.

Virtualized Desktops: 2011 promises to be a year of increased virtualized desktop adoption. Frustrated with enterprise desktop application licensing, plus desktop support model, IT business leaders will turn toward virtualizing desktops at increasing numbers. The application model of virtualized desktops is to deliver a wide range of corporate applications hosted in private/public data center clouds over the enterprise network. While there are no estimates to the traffic load this will place on campus and data center Ethernet networking, one can only assume it will add north-south traffic volume to an already increasing and dominating east-west flows.

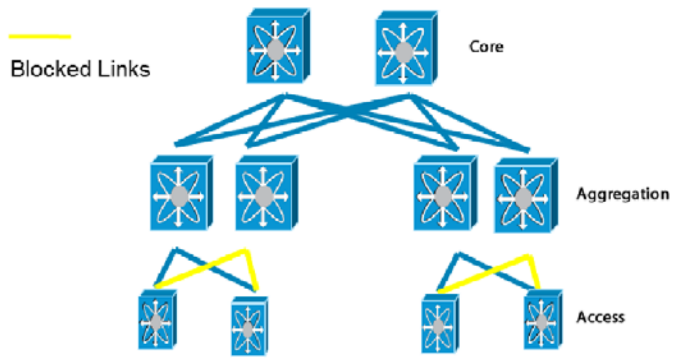
Fewer Network Tiers: To deliver low latency and high throughput performance to support increasing and changing traffic profiles mentioned above, plus ease VM moves without network configuration changes, a new two-tier private/public data center cloud network design has emerged. A three-tier network architecture is the dominant structure in data centers today and will likely continue as the optimal design for many smaller data center networks. For most network architects and administrators, this type of design provides the best balance of asset utilization, layer 3 routing for segmentation, scaling and services plus efficient physical design for cabling and fiber runs.

By three tiers, we mean access switches/ToR switches, or modular/End-of-Row switches that connect to servers and IP-based storage. These access switches are connected via Ethernet to aggregation switches. The aggregation switches are connected into a set of Core switches or routers that forward traffic flows from servers to an intranet and internet, and between the aggregation switches. It is common in this structure to oversubscribe bandwidth in the access tier, and to a lesser degree, in the aggregation tier, which can increase latency and reduce performance. Spanning Tree Protocol or STP between access and aggregation plus aggregation and core further drive oversubscription. Inherent in this structure is the placement of layer 2 versus layer 3 forwarding that is VLANs and IP routing. It is common that VLANs are constructed within access and aggregation switches, while layer 3 capabilities in the aggregation or Core switches route between them.

Within the private and public data center cloud network market, where the number of servers are in the thousands to tens of thousands plus, east-west bandwidth is significant and applications require a single layer 2 domain, the existing Ethernet or layer 2 capabilities within this tiered architecture do not meet emerging demands.

© Lippis Enterprises, Inc. 2011

Traditional three-tier Network Fabric Architecture



Two-Tier Leaf-Spine Network Fabric Architecture

An increasingly common design to scale a data center fabric is often called a "fat-tree" or "leaf-spine" and consists of two kinds of switches; one that connects servers and the second that connects switches creating a non-blocking, low-latency fabric. We use the terms "leaf" switch to denote server connecting switches and "spine" to denote switches that connect leaf switches. Together, a leaf and spine architecture create a scalable data center fabric. Key to this design is the elimination of STP with some number of multi-links between leaf and spine that eliminate oversubscription and enable a non-blocking fabric, assuming the switches are designed with enough backplane capacity to support packet forwarding equal to the sum of leaf ingress bandwidth. There are multiple approaches for connecting leaf and spine switches at high bandwidth, which fall under the category of Multi-Chassis Link Aggregation Group or MC-LAG covered in project IEEE 802.3ad.

Paramount in the two-tier leaf-spine architecture is high spine switch performance, which collapses the aggregation layer in the traditional three-tier network. Another design is to connect every switch together in a full mesh, via MC-LAG connections with every server being one hop away from each other.

The above captures the major trends and demands IT business leaders are requiring from the networking industry. The underpinnings of private and public data center cloud network fabric are 10GbE switching with 40GbE and 100GbE ports/modules. 40GbE and 100GbE are in limited availability now but will be increasingly offered and adopted during 2011. Network performance including throughput performance and latency are fundamental switch attributes to understand and review across suppliers. Because if the 10GbE switches an IT leader selects cannot scale performance to support increasing traffic volume plus shifts in traffic profile, not only will the network fail to be a fabric unable to support converge storage traffic, but business processes, application performance and user experience will suffer too.

During 2011, an increasing number of servers will be equipped with 10GbE LAN on Motherboard (LOM) driving 10GbE network requirements. In addition, with nearly 80% of IT spend being consumed in data center infrastructure with all IT assets eventually running over 10GbE switching, the stakes could not be higher to select the right product upon which to build this fundamental corporate asset. Further, data center network equipment has the longest life span of all IT equipment; therefore, networking is a long-term investment and vendor commitment.

© Lippis Enterprises, Inc. 2011

The Lippis Report Test Methodology

To test products, each supplier brought its engineers to configure its equipment for test. An Ixia test engineer was available to assist each supplier through test methodologies and review test data. After testing was concluded, each supplier's engineer signed off on the resulting test data. We call the following set of testing conducted "The Lippis Test." The test methodologies included:

Throughput Performance: Throughput, packet loss and delay for layer-2 (L2) unicast, layer-3 (L3) unicast and layer-3 multicast traffic was measured for packet sizes of 64, 128, 256, 512, 1024, 1280, 1518, 2176, 9216 bytes. In addition, a special cloud computing simulation throughput test consisting of a mix of north-south plus east-west traffic was conducted. Ixia's IxNetwork RFC 2544 Throughput/Latency quick test was used to perform all but the multicast tests. Ixia's IxAutomate RFC 3918 Throughput No Drop Rate test was used for the multicast test.

Latency: Latency was measured for all the above packet sizes plus the special mix of north-south and east-west traffic blend. Two latency tests were conducted: 1) latency was measured as packets flow between two ports on different modules for modular switches, and 2) between far away ports (port pairing) for ToR switches to demonstrate latency consistency across the forwarding engine chip. Latency test port configuration was via port pairing across the entire device versus side-by-side. This meant that a switch with N ports, port 1 was paired with port (N/2)+1, port 2 with port (N/2)+2, etc. Ixia's IxNetwork RFC 2544 Throughput / Latency quick test was used for validation.

Jitter: Jitter statistics was measured during the above throughput and latency test using Ixia's IxNetwork RFC 2544 Throughput/Latency quick test.

Congestion Control Test: Ixia's IxNetwork RFC 2889 Congestion test was used to test both L2 and L3 packets. The objective of the Congestion Control Test is to determine how a Device Under Test (DUT) handles congestion. Does the device implement congestion control and does congestion on one port affect an uncongested port? This procedure determines if Head-of-Line (HOL) blocking and/or if back pressure are present. If there is frame loss at the uncongested port, HOL blocking is present. Therefore, the DUT cannot forward the amount of traffic to the congested port, and as a result, it is also losing frames destined to the uncongested port. If there is no frame loss on the congested port and the port receives more packets than the maximum offered load of 100%, then Back Pressure is present.

Video feature: **Click to view a discussion on the Lippis Report Test Methodology**

RFC 2544 Throughput/Latency Test

Test Objective: This test determines the processing overhead of the DUT required to forward frames and the maximum rate of receiving and forwarding frames without frame loss.

Test Methodology: The test starts by sending frames at a specified rate, usually the maximum theoretical rate of the port while frame loss is monitored. Frames are sent from and received at all ports on the DUT, and the transmission and reception rates are recorded. A binary, step or combo search algorithm is used to identify the maximum rate at which no frame loss is experienced.

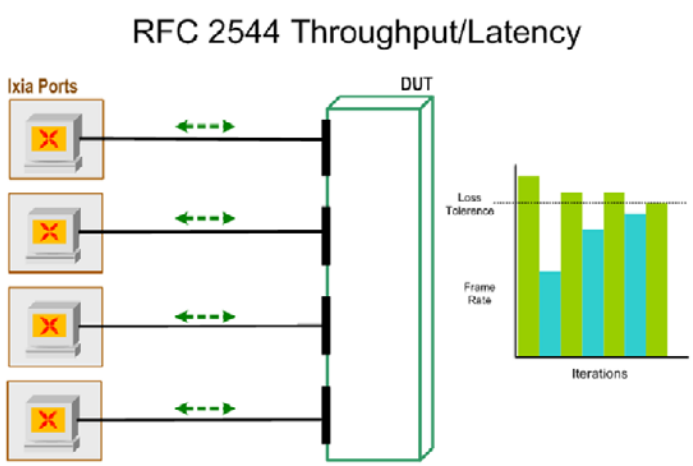
To determine latency, frames are transmitted for a fixed duration. Frames are tagged once in each second and during half of the transmission duration, then tagged frames are transmitted. The receiving and transmitting timestamp on the tagged frames are compared. The difference between

© Lippis Enterprises, Inc. 2011

the two timestamps is the latency. The test uses a one-to-one traffic mapping. For store and forward DUT switches latency is defined in RFC 1242 as the time interval starting when the last bit of the input frame reaches the input port and ending when the first bit of the output frame is seen on the output port. Thus latency is not dependent on link speed only, but processing time too.

Results: This test captures the following data: total number of frames transmitted from all ports, total number of frames received on all ports, percentage of lost frames for each frame size plus latency, jitter, sequence errors and data integrity error.

The following graphic depicts the RFC 2554 throughput performance and latency test conducted at the iSimCity lab for each product.



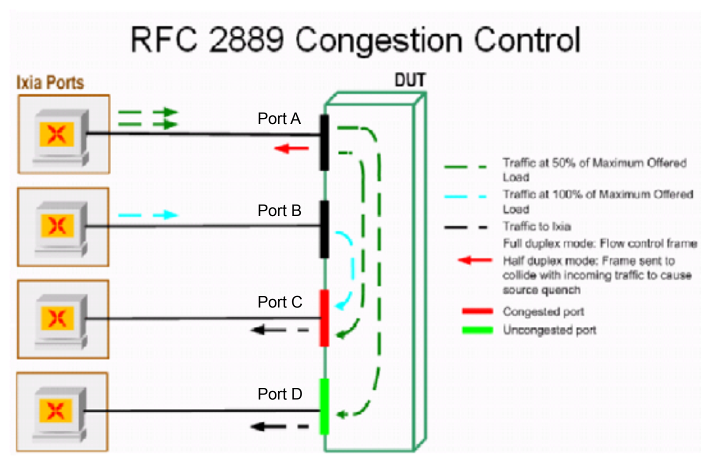
RFC 2889 Congestion Control Test

Test Objective: The objective of the Congestion Control Test is to determine how a DUT handles congestion. Does the device implement congestion control and does congestion on one port affect an uncongested port? This procedure determines if HOL blocking and/or if back pressure are present. If there is frame loss at the uncongested port, HOL blocking is present. If the DUT cannot forward the amount of traffic to the congested port, and as a result, it is also losing frames destined to the uncongested port. If there is no frame loss on the congested port and the port receives more packets than the maximum offered load of 100%, then Back Pressure is present.

Test Methodology: If the ports are set to half duplex, collisions should be detected on the transmitting interfaces. If the ports are set to full duplex and flow control is enabled, flow control frames should be detected. This test consists of a multiple of four ports with the same MOL (Maximum Offered Load). The custom port group mapping is formed of two ports, A and B, transmitting to a third port C (the congested interface), while port A also transmits to port D (uncongested interface).

Test Results: This test captures the following data: intended load, offered load, number of transmitted frames, number of received frames, frame loss, number of collisions and number of flow control frames obtained for each frame size of each trial are captured and calculated.

The following graphic depicts the RFC 2889 Congestion Control test as conducted at the iSimCity lab for each product.



RFC 3918 IP Multicast Throughput No Drop Rate Test

Test Objective: This test determines the maximum throughput the DUT can support while receiving and transmitting multicast traffic. The input includes protocol parameters (IGMP, PIM), receiver parameters (group addressing), source parameters (emulated PIM routers), frame sizes, initial line rate and search type.

Test Methodology: This test calculates the maximum DUT throughput for IP Multicast traffic using either a binary or a linear search, and to collect Latency and Data

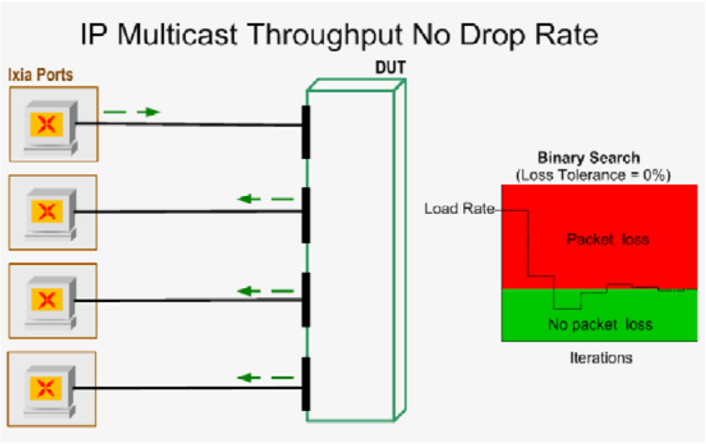
 © Lippis Enterprises, Inc. 2011

Integrity statistics. The test is patterned after the ATSS Throughput test; however this test uses multicast traffic. A one-to-many traffic mapping is used, with a minimum of two ports required.

If choosing OSPF or ISIS as IGP protocol routing, the transmit port first establishes an IGP routing protocol session and PIM session with the DUT. IGMP joins are then established for each group, on each receive port. Once protocol sessions are established, traffic begins to transmit into the DUT and a binary or linear search for maximum throughput begins.

If choosing "none" as IGP protocol routing, the transmit port does not emulate routers and does not export routes to virtual sources. The source addresses are the IP addresses configured on the Tx ports in data frame. Once the routes are configured, traffic begins to transmit into the DUT and a binary or linear search for maximum throughput begins.

Test Results: This test captures the following data: maximum throughput per port, frame loss per multicast group, minimum/maximum/average latency per multicast group and data errors per port.The following graphic depicts the RFC 3918 IP Multicast Throughput No Drop Rate test as conducted at the iSimCity lab for each product.



Power Consumption Test

Port Power Consumption: Ixia's IxGreen within the IxAutomate test suite was used to test power consumption at the port level under various loads or line rates.

Test Objective: This test determines the Energy Consumption Ratio (ECR), the ATIS (Alliance for Telecommunications Industry Solutions) Telecommunications Energy Efficiency Ratio (TEER) during a L2/L3 forwarding performance. TEER is a measure of network-element efficiency quantifying a network component's ratio of "work performed" to energy consumed.

Test Methodology: This test performs a calibration test to determine the no loss throughput of the DUT. Once the maximum throughput is determined, the test runs in automatic or manual mode to determine the L2/L3 forwarding performance while concurrently making power, current and voltage readings from the power device. Upon completion of the test, the data plane performance and Green (ECR and TEER) measurements are calculated. Engineers followed the methodology prescribed by two ATIS standards documents:

ATIS-0600015.03.2009: Energy Efficiency for Telecommunication Equipment: Methodology for Measuring and Reporting for Router and Ethernet Switch Products, and

ATIS-0600015.2009: Energy Efficiency for Telecommunication Equipment: Methodology for Measuring and Reporting - General Requirements

The power consumption of each product was measured at various load points: idle 0%, 30% and 100%. The final power consumption was reported as a weighted average calculated using the formula:

WATIS = 0.1*(Power draw at 0% load) + 0.8*(Power draw at 30% load) + 0.1*(Power draw at 100% load).

All measurements were taken over a period of 60 seconds at each load level, and repeated three times to ensure result repeatability. The final WATIS results were reported as a weighted average divided by the total number of ports per switch to derive at a WATTS per port measured per ATIS methodology and labeled here as $WATTS_{ATIS.}$

Test Results: The L2/L3 performance results include a measurement of WATIS and the DUT TEER value. Note that a larger TEER value is better as it represents more work done at less energy consumption. In the graphics

© Lippis Enterprises, Inc. 2011

throughout this report, we use $WATTS_{ATIS}$ to identify ATIS power consumption measurement on a per port basis.

With the $WATTS_{ATIS}$ we calculate a three-year energy cost based upon the following formula.

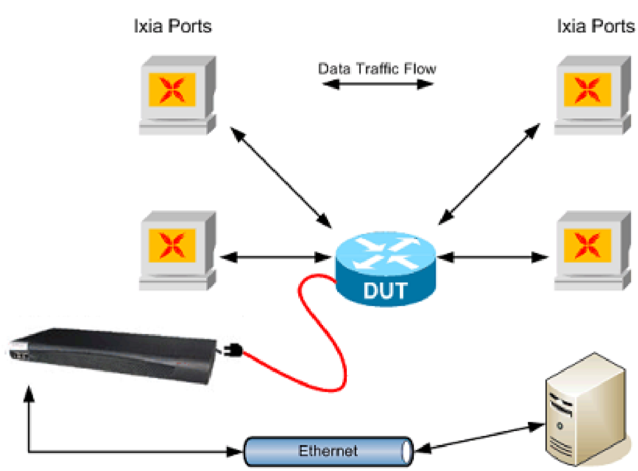
Cost/$Watts_{ATIS}$/3-Year =
($WATTS_{ATIS}$ /1000)*(3*365*24)*(0.1046)*(1.33),
where $WATTS_{ATIS}$ = ATIS weighted average power in Watts

3*365*24 = 3 years @ 365 days/yr @ 24 hrs/day

0.1046 = U.S. average retail cost (in US$) of commercial grade power as of June 2010 as per Dept. of Energy Electric Power Monthly

(http://www.eia.doe.gov/cneaf/electricity/epm/table5_6_a.html)
1.33 = Factor to account for power costs plus cooling costs @ 33% of power costs.

The following graphic depicts the per port power consumption test as conducted at the iSimCity lab for each product.



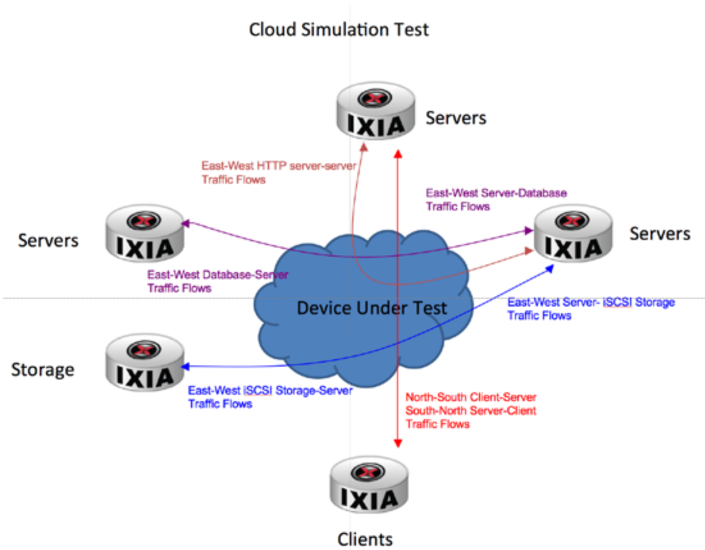
Public Cloud Simulation Test

Test Objective: This test determines the traffic delivery performance of the DUT in forwarding a variety of north-south and east-west traffic in cloud computing applications. The input parameters include traffic types, traffic rate, frame sizes, offered traffic behavior and traffic mesh.

Test Methodology: This test measures the throughput, latency, jitter and loss on a per application traffic type basis across M sets of 8-port topologies. M is an integer and is proportional to the number of ports the DUT is populated with. This test includes a mix of north-south traffic and east-west traffic, and each traffic type is configured for the following parameters: frame rate, frame size distribution, offered traffic load and traffic mesh. The following traffic types are used: web (HTTP), database-server, server-database, iSCSI storage-server, iSCSI server-storage, client-server plus server-client. The north-south client-server traffic simulates Internet browsing, the database traffic simulates server-server lookup and data retrieval, while the storage traffic simulates IP-based storage requests and retrieval. When all traffic is transmitted, the throughput, latency, jitter and loss performance are measured on a per traffic type basis.

Test Results: This test captures the following data: maximum throughput per traffic type, frame loss per traffic type, minimum/maximum/average latency per traffic type, minimum/maximum/average jitter per traffic type, data integrity errors per port and CRC errors per port. For this report we show average latency on a per traffic basis at zero frame loss.

The following graphic depicts the Cloud Simulation test as conducted at the iSimCity lab for each product.



© Lippis Enterprises, Inc. 2011

Participating Suppliers

This was the first time that all participating vendors' switches were tested in an open public forum. In addition, the ToR switches all utilized a new single chip design sourced from Broadcom, Marvel or Fulcrum. This chip design supports up to 64 10GbE ports on a single chip and represents a new generation of ToR switches. These ToR or leaf switches with price per port as low as $350 per 10GbE are the fastest growing segment of the 10GbE fixed configuration switch market. Their list price points vary $12K to $25K. ToR switches tested were:

Arista 7124SX 10G SFP Data Center Switch

Arista 7050S-64 10/40G Data Center Switch

BLADE Network Technologies, an IBM Company
IBM BNT RackSwitch G8124

BLADE Network Technologies, an IBM Company
IBM BNT RackSwitch G8264

Brocade VDX™ 6720-24 Data Center Switch

Force10 S-Series S4810

Hitachi Cable, Apresia15000-64XL-PSR

Voltaire® Vantage™ 6048

All Core switches including Juniper's EX8216, Arista Networks' 7504, and Alcatel-Lucent's OmniSwitch 10K were tested for the first time in an open and public forum too. These "spine" switches make up the core of cloud networking. A group of spine switches connected in partial or full mesh creates an Ethernet fabric connecting ToR or leaf switches and associated data plus storage traffic. These Core switches range in list price from $250K to $800K and price per 10GbE port of $1.2K to $6K depending upon port density and software license arrangement. The Core/Spine switches tested were:

Alcatel-Lucent OmniSwitch 10K

Arista 7504 Series Data Center Switch

Juniper Network EX Series EX8200 Ethernet Switch

We present each supplier's results in alphabetical order.

© Lippis Enterprises, Inc. 2011

Alcatel-Lucent OmniSwitch™ 10K

Alcatel-Lucent launched its new entry into the enterprise data center market on December 17, 2010, with the OmniSwitch™ 10K. The OmniSwitch was the most densely-populated device tested with 256 ports of 10GbE. The test numbers below represent the first public performance and power consumption measurements for the OmniSwitch™ 10K. The Alcatel-Lucent OmniSwitch™ 10K Modular Ethernet LAN Chassis is the first of a new generation of network adaptable LAN switches. It exemplifies Alcatel-Lucent's approach to enabling what it calls Application Fluent Networks, which are designed to deliver a high-quality user experience while optimizing the performance of legacy, real-time, and multimedia applications.

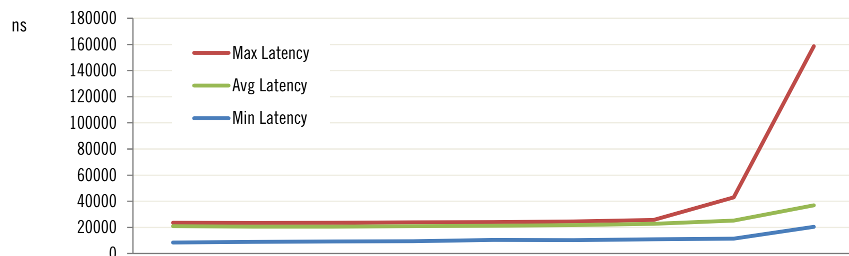
Video feature: Click to view Alcatel-Lucent video podcast

The OmniSwitch 10K was tested across all 256 ports of 10GbE. Its average latency ranged from a low of 20561 ns or 20 **µs** to a high of 36,823 ns or 36 µs at jumbo size 9216 Byte size frames for layer 2 traffic. Its average delay variation was ranged between 5 and 10 ns, providing consistent latency across all packet sizes at full line rate.

Alcatel-Lucent OmniSwitch 10K Test Configuration

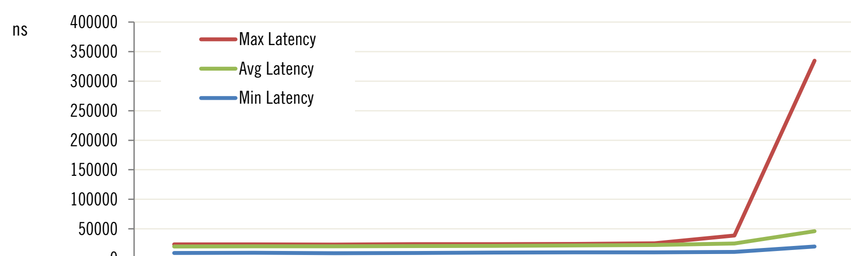
	Hardware	Software Version	Port Density
Device under test	OmniSwitch 10K www.alcatel-lucent.com	7.1.1.R01.1638	256
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Alcatel-Lucent OmniSwitch 10K* RFC 2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	23600	23420	23580	23840	24040	24480	25640	42960	158600
Avg Latency	20864	20561	20631	20936	21216	21787	22668	25255	36823
Min Latency	8480	8940	9280	9500	10380	10300	10960	11440	20420
Avg. delay Variation	9	5	5	8	7	5	10	5	10

Alcatel-Lucent OmniSwitch 10K* RFC 2544 Layer 3 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	23700	23600	23500	24000	24280	24480	25700	38580	334680
Avg Latency	20128	20567	20638	20931	21211	21793	22658	25242	45933
Min Latency	9160	9360	8760	9260	9740	10340	10260	10980	20180
Avg. delay Variation	9	5	5	8	7	4	10	5	10

* These RFC 2544 tests were conducted with default configuration. Lower latency configuration mode will be available in the future.

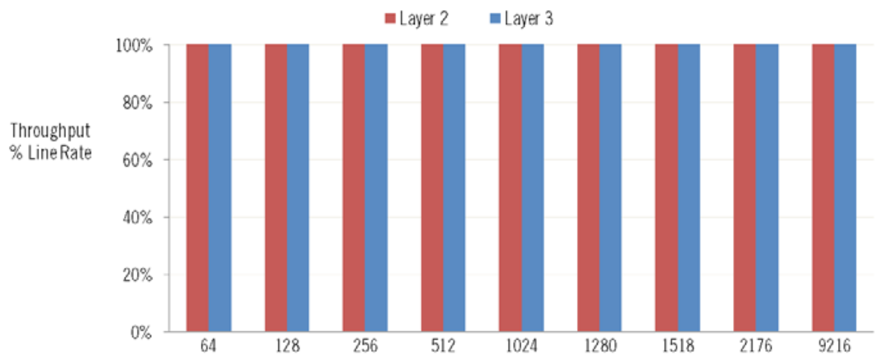
© Lippis Enterprises, Inc. 2011

For layer 3 traffic, the OmniSwitch 10K's measured average latency ranged from a low of 20,128 ns at 64Bytes to a high of 45,933 ns or 45μs at jumbo size 9216 Byte size frames. Its average delay variation for layer 3 traffic ranged between 4 and 10 ns, providing consistent latency across all packet sizes at full line rate.

The OmniSwitch 10K demonstrated 100% throughput as a percentage of line rate across all 256-10GbE ports. In other words, not a single packet was dropped while the OmniSwitch 10K was presented with enough traffic to populate its 256 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

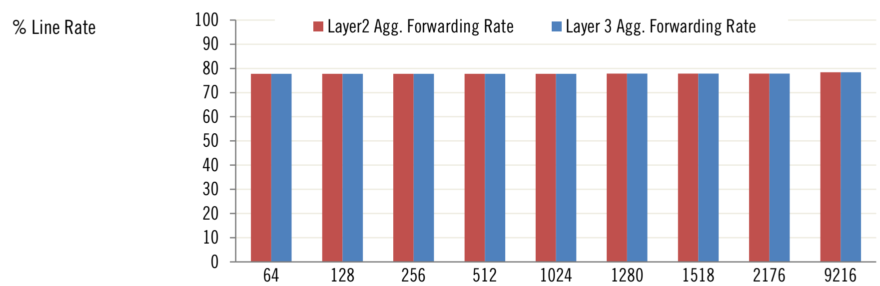
The OmniSwitch 10K demonstrated nearly 80% of aggregated forwarding rate as percentage of line rate during congestion conditions. A single 10GbE port was flooded at 150% of line rate. The OmniSwitch did not use HOL blocking which means that as the 10GbE port on the OmniSwitch became congested, it did not impact the performance of other ports. There was no back pressure detected, and the Ixia test gear did not receive flow control frames.

Alcatel-Lucent OmniSwitch 10K RFC 2544 L2 & L3 Throughput Test

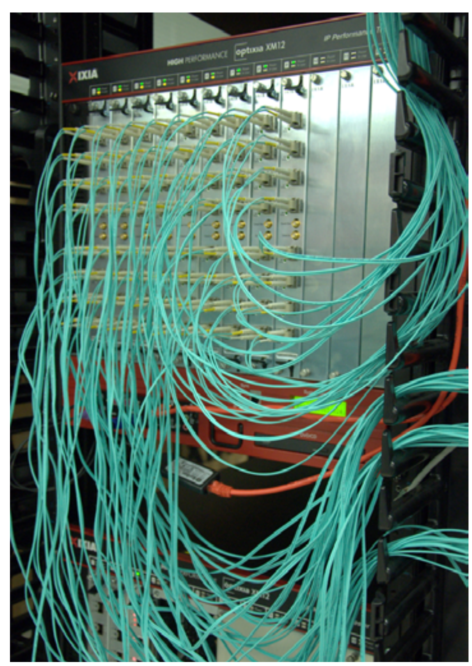


	64	128	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100

Alcatel-Lucent OmniSwitch 10K RFC 2889 Congestion Test

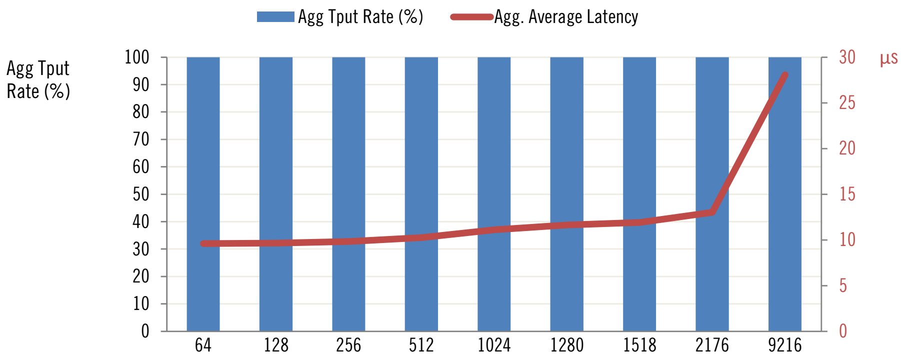


	64	128	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	77.784	77.79	77.786	77.772	77.775	77.813	77.828	77.857	78.39
Layer 3 Agg. Forwarding Rate	77.784	77.79	77.786	77.772	77.775	77.813	77.828	77.856	78.39
Head of Line Blocking	no	no	no	no	no	no	no	no	no
Back Pressure	no	no	no	no	no	no	no	no	no
Agg Flow Control Frames	0	0	0	0	0	0	0	0	0



© Lippis Enterprises, Inc. 2011

Alcatel-Lucent OmniSwitch 10K RFC 3918 IP Multicast Test



	64	128	256	512	1024	1280	1518	2176	9216
Agg Tput Rate (%)	100	100	100	100	100	100	100	100	100
Agg. Average Latency	9.6	9.6	9.8	10.3	11.1	11.6	11.9	13.0	28.1

The OmniSwitch 10K demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a 9,596 ns at 64Byte size packet to 28,059 ns at 9216Byte size packets.

Alcatel-Lucent OmniSwitch10K Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	28125
East-West	Server_to_Database	14063
East-West	HTTP	19564
East-West	iSCSI-Server_to_Storage	17140
East-West	iSCSI-Storage_to_Server	26780
North-South	Client_to_Server	13194
North-South	Server_to_Client	11225

The OmniSwitch 10K performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 28μs.

Alcatel-Lucent OmniSwitch10K Power Consumption Test

$Watts_{ATIS}$/10GbE port	13.34
3-Year Cost/$Watts_{ATIS}$/10GbE	$48.77
Total power cost/3-Year	$12.485.46
3 yr energy cost as a % of list price	2.21%
TEER Value	71
Cooling	Front to Back

The OmniSwitch 10K represents a new breed of cloud network spine switches with power efficiency being a core value. Its $Watts_{ATIS}$/port is 13.34 and TEER value is 71. Its power cost per 10GbE is estimated at $16.26 per year. The three-year cost to power the OmniSwitch is estimated at $12,485.46 and represents less than 3% of its list price. Keeping with data center best practices, its cooling fans flow air front to back.

© Lippis Enterprises, Inc. 2011

Discussion:

The OmniSwitch™ 10K seeks to improve application performance and user experience with deep packet buffers, lossless virtual output queuing (VOQ) fabric and extensive traffic management capabilities. This architecture proved its value during the RFC2889 layer 2 and layer 3 congestion test with a 78% aggregated forwarding rate when a single 10GbE port was oversubscribed at 150% of line rate. The OmniSwitch™ 10K did not use HOL blocking, back pressure or signal back to the Ixia test equipment with Aggregated Flow Control Frames to slow down traffic flow. Not tested but notable features are its security and high availability design for uninterrupted uptime.

The OmniSwitch™ 10K was found to have low power consumption, front-to-back cooling, front-accessible components and a compact form factor. The OmniSwitch™ 10K is designed to meet the requirements for mid- to large-sized enterprises data centers.

© Lippis Enterprises, Inc. 2011

Arista 7504 Series Data Center Switch

On January 8, 2011, Arista Networks launched a new member of its 7500 series family called the 7504 modular switch that supports a 192 10GbE interface. The test data and graphics below represent the first public performance and power consumption measurements of the 7504. The 7504 is a smaller version of its award-winning and larger 384 10GbE port 7508 data center switch.

Video feature: Click to view Arista video podcast

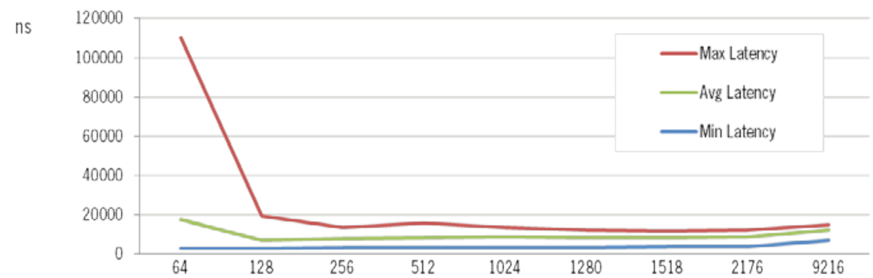
The Arista 7504 was tested across all 192 ports of 10GbE. Its average latency ranged from a low of 6,831 ns or 6.8 **µs** to a high of 17,791 ns or 17.7 µs at jumbo size 64 Byte size frames for layer 2 traffic. Its average delay variation was ranged between 5 and 10 ns, providing consistent latency across all packet sizes at full line rate.

For layer 3 traffic, the Arista 7504's average latency ranged from a low of 6,821 ns to a high of 12,387 ns or 12.3 µs at jumbo size 9216 Byte size frames. Its average delay variation for layer 3 traffic ranged between 5 and 10.6 ns, providing consistent latency across all packet sizes at full line rate.

Arista Networks 7504 Test Configuration

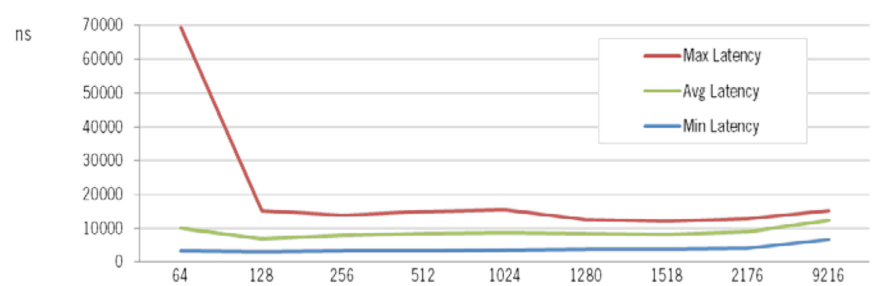
	Hardware	Software Version	Port Density
Device under test	7504 Modular Switch http://www.aristanetworks.com	EOS 4.5.5	192
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Arista 7504 RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	110120	19560	13800	16020	13540	12520	12040	12500	15160
Avg Latency	17791.203	6831.521	7850.714	8452.021	8635.505	8387.01	8122.078	8995.995	12392.828
Min Latency	3080	3100	3280	3420	3580	3640	3720	3960	6840
Avg. delay Variation	9	5	5.953	8	7	6	10	5	10.547

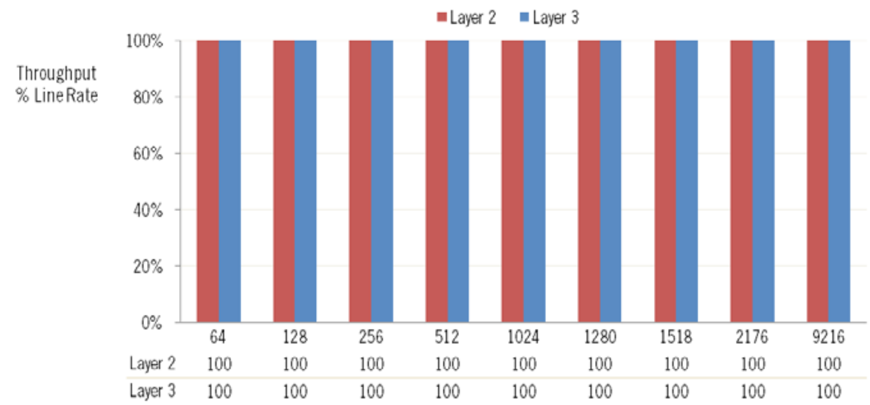
Arista 7504 RFC2544 Layer 3 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	69420	15300	13880	15060	15340	12720	12060	12820	15160
Avg Latency	9994.88	6821.323	7858.688	8466.526	8647.958	8387.26	8112.948	8996.547	12387.99
Min Latency	3340	3160	3440	3380	3660	3740	3720	3980	6600
Avg. delay Variation	9	5	5.953	8	7	6	10	5	10.599

© Lippis Enterprises, Inc. 2011

Arista 7504
RFC 2544 L2 & L3 Throughput Test



The Arista 7504 demonstrated 100% throughput as a percentage of line rate across all 192 10GbE ports. In other words, not a single packet was dropped while the 7504 was presented with enough traffic to populate its 192 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

Arista 7504 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	14208
East-West	Server_to_Database	4394
East-West	HTTP	9171
East-West	iSCSI-Server_to_Storage	5656
East-West	iSCSI-Storage_to_Server	13245
North-South	Client_to_Server	5428
North-South	Server_to_Client	4225

The Arista 7504 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 14µs.

Arista 7504 Power Consumption Test

Watts_{ATIS}/10GbE port	10.31
3-Year Cost/Watts_{ATIS}/10GbE	$37.69
Total power cost/3-Year	$7,237.17
3 yr energy cost as a % of list price	3.14%
TEER Value	92
Cooling	Front to Back

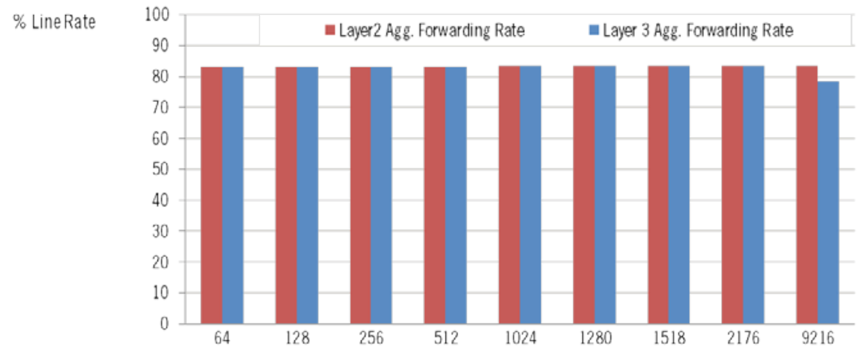
The Arista 7504 represents a new breed of cloud network spine switches with power efficiency being a core value. Its Watts_{ATIS}/port is 10.31 and TEER value is 92. Its power cost per 10GbE is estimated at $12.56 per year. The three-year cost to power the 7504 is estimated at $7,237.17 and represents approximately 3% of its list price. Keeping with data center best practices, its cooling fans flow air front to back.

© Lippis Enterprises, Inc. 2011

Arista 7504
RFC 3918 IP Multicast Test

The Arista 7504 does not support IP multicast at this time.

Arista 7504
RFC 2889 Congestion Test



	64	128	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	83.34	83.342	83.345	83.35	83.362	83.362	83.362	83.362	83.362
Layer 3 Agg. Forwarding Rate	83.34	83.342	83.345	83.35	83.362	83.362	83.362	83.362	78.209
L2 Head of Line Blocking	no	no	no	no	no	no	no	no	no
L3 Head of Line Blocking	no	no	no	no	no	no	no	no	yes
L2 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes
L3 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	no
Agg Flow Control Frames	0	0	0	0	0	0	0	0	0

The Arista 7504 demonstrated nearly 84% of aggregated forwarding rate as percentage of line rate during congestion conditions for both L2 and L3 traffic flows. A single 10GbE port was flooded at 150% of line rate. For L2 forwarding the 7504 did not use HOL blocking, which means that as the 7504-10GbE port became congested, it did not impact the performance of other ports. However, during L3 testing HOL blocking was detected at the 9216 packet size. Note the 7504 was in beta testing at the time of the Lippis/IXIA test and there was a "*corner case*" at 9216 bytes at L3 that needed further tuning. Arista states that its production code provides wirespeed L2/L3 performance at all packet sizes without any head of line blocking.

Note that while the Arista 7504 shows back pressure, in fact there is none. IXIA and other test equipment calculates back pressure per RFC 2889 paragraph 5.5.5.2. which states that if the total number of received frames on the congestion port surpasses the number of transmitted frames at MOL (Maximum Offered Load) rate then back pressure is present. Thanks to the 7504's 2.3GB of packet buffer memory it can overload ports with more packets than the MOL, therefore, the IXIA or any test equipment "calculates/sees" back pressure, but in reality this is an anomaly of the RFC testing method and not the 7504. The Arista 7504 can buffer up 40ms of traffic per port at 10GbE speeds which is 400K bits or 5,425 packets of 9216 bytes.

Arista 7500 switches do not use back pressure (PAUSE) to manage congestion, while other core switches in the industry, not tested here, transmit pause frames under congestion. The Arista 7504 offers lossless forwarding for datacenter applications due to its deep buffer VOQ architecture. With 2.3GB of packet buffer memory, Arista could not find any use case where the 7504 needs to transmit pause frames and hence that functionality is turned off.

The Arista 7504 does honor pause frames and stops transmitting, as many hosts and other switches still cannot support wirespeed 10G traffic. Again due to its deep buffers, the 7504 can buffer up to 40ms of traffic per port at 10GbE speeds, which is a necessary factor for a lossless network.

Arista's excellent congestion management results are a direct consequence of its generous buffer memory design as well as its VOQ architecture.

© Lippis Enterprises, Inc. 2011

Discussion:

In a compact 7 RU chassis, the Arista 7504 offers 5 Terabits of fabric capacity, 192 wire speed 10 GbE ports and 2.88 BPPS of L2/3 throughput. The Arista 7500 boasts a comprehensive feature set, as well as future support for 40GbE and 100GbE standard without needing a fabric upgrade. According to Arista with front-to-rear airflow, redundant and hot swappable supervisor, power, fabric and cooling modules, the 7500 is energy efficient with typical power consumption of approximately 10 watts per port for a fully loaded chassis. In fact, the Lippis/Ixia test measures 10.31 $WATTS_{ATIS}$ per port for the 7504 with a TEER value of 92. The Arista 7500 is a platform targeted for building low latency, scalable, data center networks.

Arista EOS is a modular switch operating system with a unique state sharing architecture that separates switch state from protocol processing and application logic. Built on top of a standard Linux kernel, all EOS processes run in their own protected memory space and exchange state through a memory-based database. This multi-process state sharing architecture provides the foundation for in-service-software updates and self-healing resiliency. Arista EOS automates the networking infrastructure natively with VMware vSphere via its VM Tracer to provide VM discovery, auto-VLAN provisioning and visibility into the virtual computing environment. All Arista products, including the 7504 switch, run Arista EOS software. The same binary image supports all Arista products.

For the Arista 7500 series, engineers have implemented an advanced VOQ architecture that eliminates HOL blocking with a dedicated queue on the ingress side for each output port and each priority level. This is verified in the RFC 2889 layer 2 and 3 congestion Lippis/Ixia test where there is no HOL blocking even when an egress 10GbE port is saturated with 150% of line rate traffic and in the RFC2544 latency Lippis/Ixia test where the average delay variation range is 5 to 10 ns across all packet sizes. As a result of VOQ, network traffic flows through the switch smoothly and without encountering additional congestion points. In addition, the Arista 7500 VOQ fabric is buffer-less. Packets are only stored once on ingress. This achieves one of the lowest store-and-forward latencies for any Core switch product, less than 10 microseconds for a 64B packet at layer 3.

© Lippis Enterprises, Inc. 2011

Arista 7124SX 10G SFP Data Center Switch

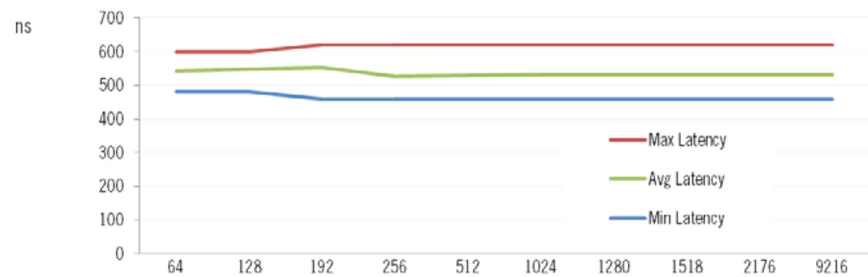
For this Lippis/Ixia evaluation at iSim-City Arista submitted its newest addition to the Arista 7100 Series Top-of-Rack (ToR) switches, the 7124SX. The 7124SX is a wire-speed 24-port cut-through switch in a 1RU (Rack Unit) form factor supporting 100Mb, 1GbE and 10GbE SPF+ optics and cables. It forwards packets at layer 2, and layer 3 in hardware. The 7124SX was built for data center and ultra low latency environments where sub 500ns latency is required. The 7124SX runs Arista's Extensible Operating System (EOS)™, which is common across the entire Arista family of switches in the same binary EOS image and can be customized to customer needs such as financial services with access to Linux tools. With EOS the 7124SX possesses in-service-software-upgrade (ISSU), self-healing stateful fault repair (SFR) and APIs to customize the software.

The 7124SX supports Arista's Latency Analyzer (LANZ), a new capability to monitor and analyze network performance, generate early warning of pending congestion events, and track sources of bottlenecks. In addition, a 50GB Solid State Disk (SSD) is available as a factory-option and can be used to store packet data captures on the switch, LANZ generated historical data, and take full advantage of the Linux based Arista EOS. Other EOS provisioning and monitoring modules include Zero Touch Provisioning (ZTP) and VM Tracer, which ease set-up and overall manageability of networks in virtualized data center environments.

Arista Networks 7124SX Test Configuration

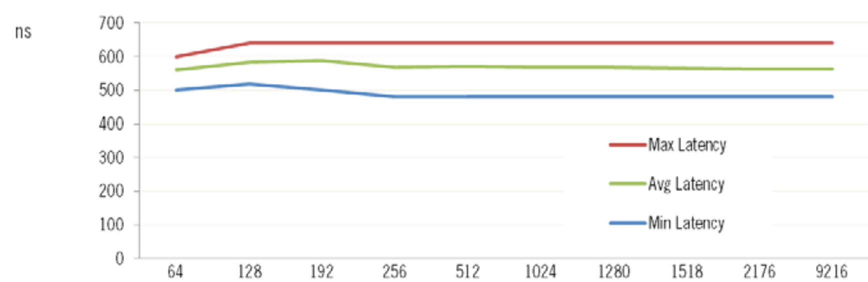
	Hardware	Software Version	Port Density
Device under test	7124SX 10G SFP Switch http://www.aristanetworks.com	EOS-4.6.2	24
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Arista 7124SX
RFC2544 Layer 2 Latency Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Max Latency	600	600	620	620	620	620	620	620	620	620
Avg Latency	541.083	548.208	553.292	527.917	528.958	531.583	531.417	531.042	531.417	532.792
Min Latency	480	480	460	460	460	460	460	460	460	460
Avg. delay Variation	8.417	5.958	9.667	5.542	7.667	7.833	5.917	9.542	6.667	9.667

Arista 7124SX
RFC2544 Layer 3 Latency Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Max Latency	600	640	640	640	640	640	640	640	640	640
Avg Latency	561.583	582.333	589.5	567.75	570.292	568.792	568.542	564.333	563.833	562.208
Min Latency	500	520	500	480	480	480	480	480	480	480
Avg. delay Variation	9.958	5.208	9.917	5.417	7.917	6.875	5.5	9.25	5.25	9.333

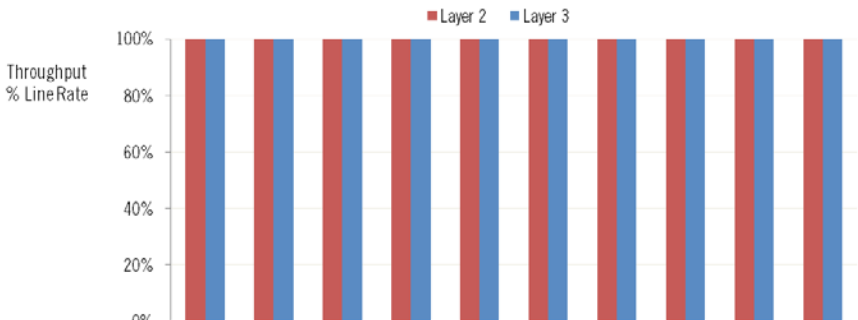
© Lippis Enterprises, Inc. 2011

For those in the financial services industry, pay particular attention to packet size 192 Bytes, which typically is not covered by many tests but is the most common packet size in financial trading environments.

The 7124SX ToR switch was tested across all 24 ports of 10GbE at layer 2 and 3 forwarding. The minimum layer 2 latency observed was 460 ns. Its average layer 2 latency was measured from a low of 527 ns to a high of 553 ns, a very tight range across packet sizes from 64 to 9216 Bytes. Its average delay variation ranged between 5.9 and 9.6 ns, providing consistent latency across all packet sizes at full line rate.

The minimum layer 3 latency measured was 480 ns. The average layer 3 latency of the 7124SX was measured from a low of 561 ns to a high of 589 ns, again a very tight range across packet sizes from 64 to 9216 Bytes. Its average delay variation ranged between 5.2 and 9.9 ns, providing consistent latency across all packet sizes at full line rate.

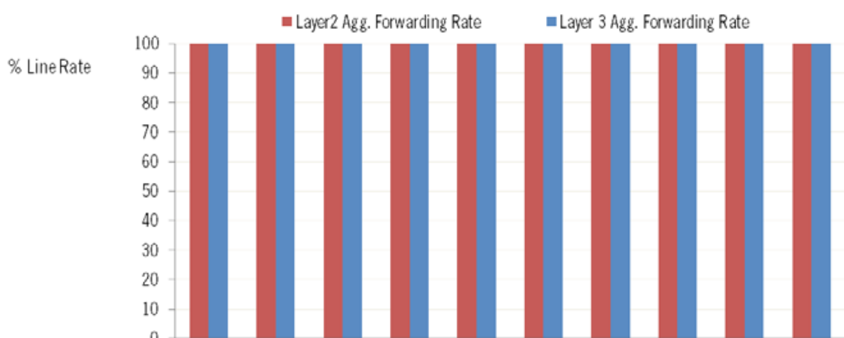
Arista 7124SX
RFC 2544 L2 & L3 Throughput Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100	100

The Arista 7124SX demonstrated 100% throughput as a percentage of line rate across all 24 10GbE ports. In other words, not a single packet was dropped while the Arista 7124SX was presented with enough traffic to populate its 24 10GbE ports at line rate simultaneously for L2 and L3 traffic flows. The 7124SX is indeed a 24-port wire-speed switch.

Arista 7124SX
RFC 2889 Congestion Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100	100
Layer 3 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100	100
L2 Head of Line Blocking	no	no	no	no	no	no	no	no	no	
L3 Head of Line Blocking	no	no	no	no	no	no	no	no	no	
L2 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes	
L3 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes	
Agg Flow Control Frames	0	0	0	0	0	0	0	0	0	

The Arista 7124SX demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions. A single 10GbE port was flooded at 150% of line rate. The Arista 7124SX did not exhibit Head of Line (HOL) blocking problems, which means that as a 10GbE port on the Arista 7124SX became congested, it did not impact performance of other ports avoiding a cascading of dropped packets and congestion.

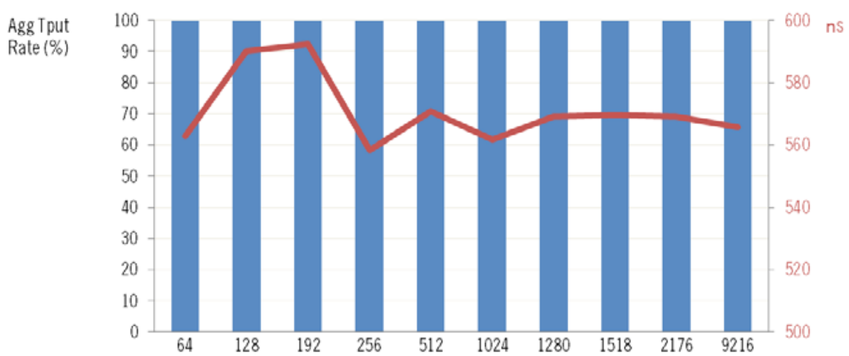
Uniquely the Arista 7124SX did indicate to Ixia test gear that it was using back pressure, however there were no flow control frames detected and in fact there were none. Ixia and other

© Lippis Enterprises, Inc. 2011

test equipment calculate back pressure per RFC 2889 paragraph 5.5.5.2., which states that if the total number of received frames on the congestion port surpasses the number of transmitted frames at MOL (Maximum Offered Load) rate then back pressure is present. Thanks to the 7124SX's generous and dynamic buffer allocation it can overload ports with more packets than the MOL, therefore, the Ixia or any test equipment "calculates/sees" back pressure, but in reality this is an anomaly of the RFC testing method and not the 7124SX.

The Arista 7124SX has a 2MB packet buffer pool and uses Dynamic Buffer Allocation (DBA) to manage congestion. Unlike other architectures that have a per-port fixed packet memory, the 7124SX can handle microbursts to allocate packet memory to the port that needs it. Such microbursts of traffic can be generated by RFC 2889 congestion tests when multiple traffic sources send traffic to the same destination for a short period of time. Using DBA, a port on the 7124SX can buffer up to 1.7MB of data in its transmit queue.

Arista 7124SX
RFC 3918 IP Multicast Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Agg Tput Rate (%)	100	100	100	100	100	100	100	100	100	100
Agg. Average Latency	562.822	590.087	592.348	558.435	570.913	561.630	569.221	569.725	569.174	565.652

The Arista 7124SX demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a 562 ns at 64Byte size packets to 565 ns at 9216 jumbo frame size packets, the lowest IP multicast latency measured thus far in the Lippis/Ixia set of tests. Further, the Arista 7124SX demonstrated consistently low IP Multicast latency.

Arista 7124SX Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	7439
East-West	Server_to_Database	692
East-West	HTTP	4679
East-West	iSCSI-Server_to_Storage	1256
East-West	iSCSI-Storage_to_Server	7685
North-South	Client_to_Server	1613
North-South	Server_to_Client	530

The Arista 7124SX performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 7685 ns even under congestion, measured in cut-through mode. Of special note is the low latency observed during north-south server-to-client and east-west server-to-database flows.

© Lippis Enterprises, Inc. 2011

Arista 7124SX Power Consumption Test	
Watts_{ATIS}/10GbE port	6.76
3-Year Cost/Watts_{ATIS}/10GbE	$24.71
Total power cost/3-Year	$593.15
3 yr energy cost as a % of list price	4.56%
TEER Value	140
Cooling	Front-to-Back, Back-to-Front

The Arista 7124SX represents a new breed of cloud network leaf or ToR switches with power efficiency being a core value. Its Watts_{ATIS}/port is a low 6.76 and TEER value is 140. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. The Arista 7124SX consumes slightly more power than a typical 100 Watt light bulb. Its power cost per 10GbE is estimated at $8.24 per year. The three-year cost to power the Arista 7124SX is estimated at $593.15 and represents approximately 4.56% of its list price. Keeping with data center best practices, its cooling fans flow air rear to front and are reversible.

© Lippis Enterprises, Inc. 2011

Discussion

The Arista 7124SX offers 24-wire speed 10GbE ports of layer 2 and 3 cut-through forwarding in a 1 RU footprint. Designed with performance in mind, the Arista 7124SX provides ultra low, 500ns latency, line-rate forwarding and unique dynamic buffer allocation to keep traffic moving under congestion and microburst conditions. These claims was verified in the Lippis/Ixia RFC2544 latency test where average latency across all packet sizes was measured between 527 to 589 ns for layer 2 and 3 forwarding. Of special note is the 7124SX's minimum latency of 460ns across a large range of packet sizes. Further, there was minimal delay variation, between 5 and 9 ns, across all packet sizes, demonstrating the switch architecture's ability to forward packets consistently and quickly boosting confidence that the Arista 7124SX will perform well in converged I/O configurations and consistently in demanding data center environments. In addition, the Arista 7124SX demonstrated 100% throughput at all packet sizes. Redundant power and fans along with various EOS high availability features ensure that the Arista 7124SX is always available and reliable for data center operations. The 7124SX is well suited for low latency applications such as financial trading, Message Passing Interface or MPI jobs, storage interconnect, virtualized environment and other 10GbE data center deployments

The 7124SX is a single chip 10GbE product, which provides consistent throughput and latency over all packet ranges as verified during the Lippis/Ixia test at iSimCity. The 7124SX is just one of the ToR switches in the Arista 7100 Series of fixed configuration ToR switches. In addition Arista also offers the 7148SX with low latency and low jitter over 48 ports and 7050S-64 ToR switch for denser 10GbE server connections that require 40GbE uplinks. These Arista ToR switch products connect into Arista's 7500 Series of modular core switches, which provide industry-leading density of 192 to 384 10GbE connections. Across all Arista switch provides is a single binary image of its EOS operation system allowing all switches to utilize and employ its high availability features as well as unique operational modules including LANZ, VM Tracer, vEOS and ZTP.

The combination of Arista's ToR and Core switches enable data center architects to build a flat, non blocking two-tier network with high 10GbE density, low latency, low power consumption and high availability. Arista's two-tier designs can scale to 18,000 nodes using standard Layer 2/Layer 3 protocols. All Arista ToR and Core switches support 32 port Multi-Chassis LAG or Link Aggregation Group which enable high levels of switch interconnect to be built creating large scale out cloud computing facilities.

© Lippis Enterprises, Inc. 2011

Arista Networks 7050S-64 Switch

For this Lippis/Ixia evaluation at iSim-City Arista submitted its highest density and newest Top-of-Rack (ToR) switch; the 7050S-64. The 7050S-64 is a wire speed layer 2/3/4 switch with 48 10GbE SFP+ and 4-40GbE QSFP+ ports in a 1RU form factor. Each 40GbE port can also operate as four independent 10GbE ports to provide a total of 64 10GbE ports. The Arista 7050S-64 forwards in cut-through or store and forward mode and boast a shared 9 MB packet buffer pool allocated dynamically to ports that are congested. The 7050S-64 runs Arista's Extensible Operating System (EOS)™, which is common across the entire Arista family of switches in the same binary EOS image. With EOS the 7050S-64 possesses in-service-software-upgrade (ISSU), self-healing stateful fault repair (SFR) and APIs to customize the software.

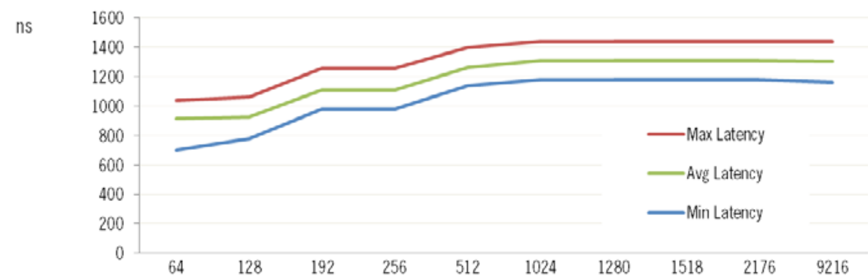
The 7050S-64 supports Zero Touch Provisioning and VMTracer to automate provisioning and monitoring of the network. In addition, a 50GB Solid State Disk (SSD) is available as a factory-option and can be used to store packet data captures on the switch, syslogs over the life of the switch, and take full advantage of the Linux based Arista EOS.

For those in the financial services industry, pay particular attention to packet size 192 Bytes, which typically is not covered by many tests but is the most common packet size in financial trading environments.

Arista Networks 7050S-64 Test Configuration

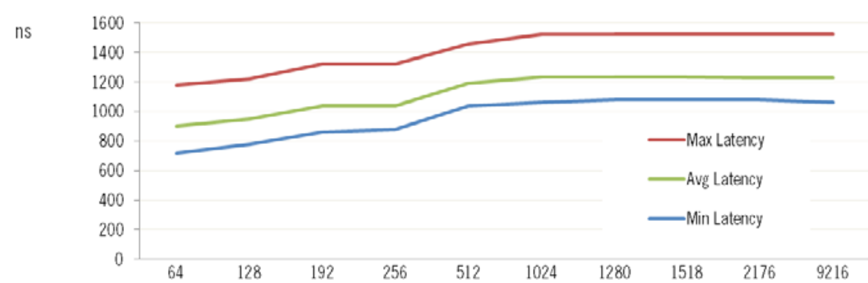
	Hardware	Software Version	Port Density
Device under test	7050S-64 Switch http://www.aristanetworks.com	EOS-4.6.2	64
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Arista 7050S-64
RFC2544 Layer 2 Latency Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Max Latency	1040	1060	1260	1260	1400	1440	1440	1440	1440	1440
Avg Latency	914.328	926.766	1110.813	1110.641	1264.906	1308.844	1309.109	1310.766	1309.766	1305.5
Min Latency	700	780	980	980	1140	1180	1180	1180	1180	1160
Avg. delay Variation	8.672	5.266	9.594	5.406	7.797	6.922	5.516	9.734	5.234	9.75

Arista 7050S-64
RFC2544 Layer 3 Latency Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Max Latency	1180	1220	1320	1320	1460	1520	1520	1520	1520	1520
Avg Latency	905.422	951.203	1038.422	1039.141	1191.641	1233	1232.453	1232.484	1229.922	1225.578
Min Latency	720	780	860	880	1040	1060	1080	1080	1080	1060
Avg. delay Variation	9.672	5.391	9.641	5.641	7.781	6.75	5.578	9.578	5.328	9.859

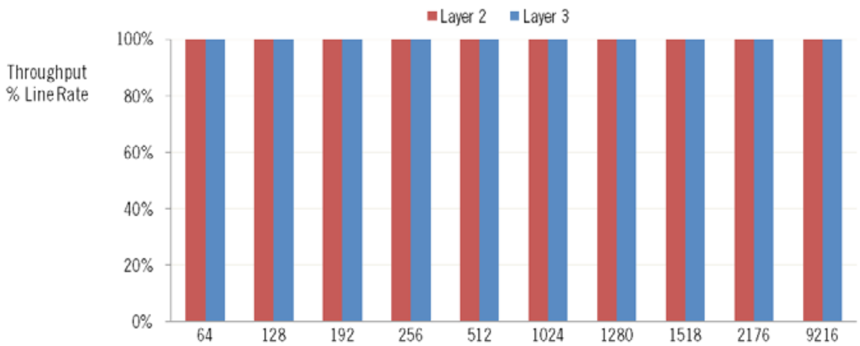
© Lippis Enterprises, Inc. 2011

The 7050S-64 ToR switch was tested in its 64-10GbE-port configuration using its 4-40GbE ports as 16-10GbE ports. It was tested for both layer 2 and 3 forwarding at 10GbE. The minimum layer 2 latency was 700 ns. The average layer 2 latency was measured from a low of 914 ns to a high of 1310 ns, across packet sizes from 64 to 9216 Bytes. Its average delay variation ranged between 5.2 and 9.7 ns, providing consistent latency across all packet sizes at full line rate.

The 7050S-64 minimum layer 3 latency was 720 ns. The 7050-64S layer 3 latency was measured from a low of 905 ns to a high of 1,233 ns, again across packet sizes from 64 to 9216 Bytes again across packet sizes from 64 to 9216 Bytes. Its average delay variation ranged between 5.3 and 9.8 ns, providing consistent latency across all packet sizes at full line rate.

The Arista 7050S-64 switch has the lowest latency amongst all 64 port switches tested so far in the Lippis/Ixia set of tests.

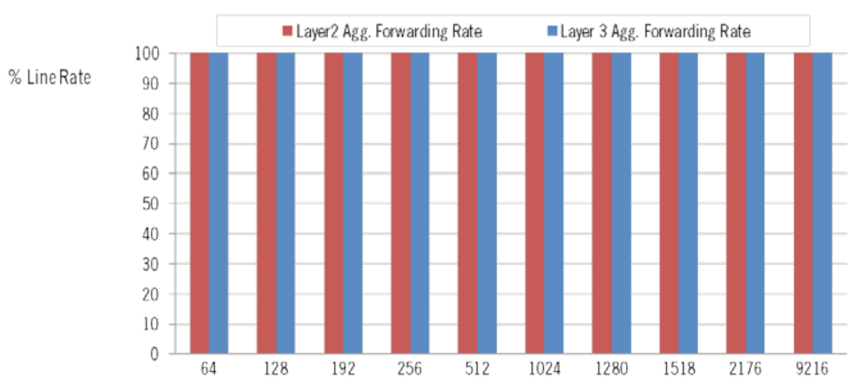
Arista 7050S-64
RFC 2544 L2 & L3 Throughput Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100	100

The Arista 7050S-64 demonstrated 100% throughput as a percentage of line rate across all 64 10GbE ports. In other words, not a single packet was dropped while the Arista 7050S-64 was presented with enough traffic to populate its 64 10GbE ports at line rate simultaneously for L2 and L3 traffic flows. The 7050S-64 is indeed a 64-port wire-speed switch.

Arista 7050S-64
RFC 2889 Congestion Test



	64	128	192	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100	100
Layer 3 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100	100
L2 Head of Line Blocking	no	no	no	no	no	no	no	no	no	no
L3 Head of Line Blocking	no	no	no	no	no	no	no	no	no	no
L2 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
L3 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Agg Flow Control Frames	0	0	0	0	0	0	0	0	0	0

The Arista 7050S-64 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions. A single 10GbE port was flooded at 150% of line rate. The Arista 7050S-64 did not exhibit Head of Line (HOL) blocking problems, which means that as a 10GbE port on the Arista 7050S-64 became congested it did not impact performance of other ports preventing a cascading of packet loss.

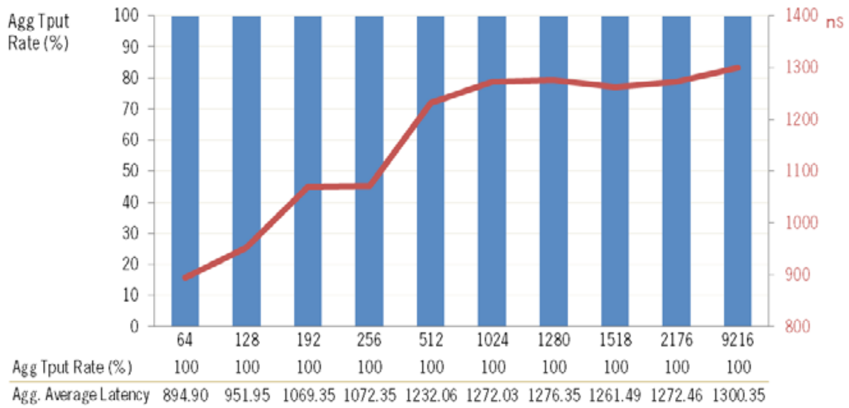
Uniquely the Arista 7050S-64 did indicate to Ixia test gear that it was using back pressure, however there were no flow control frames detected and in fact there were none. Ixia and other

© Lippis Enterprises, Inc. 2011

test equipment calculate back pressure per RFC 2889 paragraph 5.5.5.2., which states that if the total number of received frames on the congestion port surpasses the number of transmitted frames at MOL (Maximum Offered Load) rate then back pressure is present. Thanks to the 7050S-64 's generous and dynamic buffer allocation it can overload ports with more packets than the MOL, therefore, the Ixia or any test equipment "calculates/sees" back pressure, but in reality this is an anomaly of the RFC testing method and not the 7050S-64.

The Arista 7050S-64 is designed with a dynamic buffer pool allocation such that during a microburst of traffic as during RFC 2889 congestion test when multiple traffic sources are destined to the same port, packets are buffered in packet memory. Unlike other architectures that have fixed per-port packet memory, the 7050S-64 uses Dynamic Buffer Allocation (DBA) to allocate packet memory to ports that need it. Under congestion, packets are buffered in shared packet memory of 9 MBytes. The 7050S-64 uses DBA to allocate up to 5MB of packet memory to a single port for lossless forwarding as observed during this RFC 2889 congestion test.

Arista 7050S-64
RFC 3918 IP Multicast Test



The Arista 7050S-64 demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a 894 ns at 64Byte size packets to 1,300 ns at 9216 jumbo frame size packets, the lowest IP multicast latency measured thus far in the Lippis/Ixia set of test.

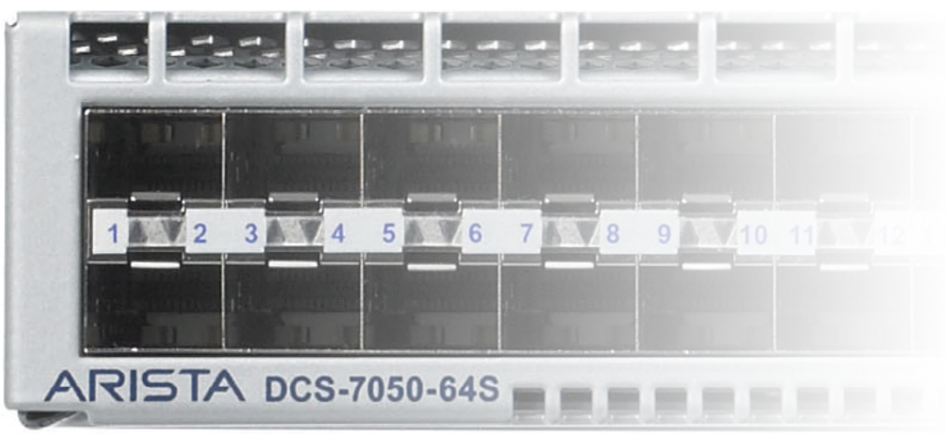
Arista 7050S-64 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	8173
East-West	Server_to_Database	948
East-West	HTTP	4609
East-West	iSCSI-Server_to_Storage	1750
East-West	iSCSI-Storage_to_Server	8101
North-South	Client_to_Server	1956
North-South	Server_to_Client	1143

The Arista 7050S-64 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 8,173 ns even under congestion, measured in cut-through mode. Of special note is the low latency observed during north-south server-to-client and east-west server-to-database flows.

© Lippis Enterprises, Inc. 2011

Arista 7050S-64 Power Consumption Test	
$Watts_{ATIS}$/10GbE port	2.34
3-Year Cost/$Watts_{ATIS}$/10GbE	$8.56
Total power cost/3-Year	$547.20
3 yr energy cost as a % of list price	1.82%
TEER Value	404
Cooling	Front-to-Back, Back-to-Front



The Arista 7050S-64 represents a new breed of cloud network leaf or ToR switches with power efficiency being a core value. Its $Watts_{ATIS}$/port is a very low 2.34 and TEER value is 404, the lowest power consumption we have measured thus far across all vendors. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. The Arista 7050S-64 consumes slightly more power than a typical 100-Watt light bulb. Its power cost per 10GbE is estimated at $2.85 per year. The three-year cost to power the Arista 7050S-64 is estimated at $547.20 and represents approximately 1.82% of its list price. Keeping with data center best practices, Arista provides both front to rear or rear to front airflow options, allowing the switch to be mounted as cabling needs dictate. Arista has taken the added step of color coding the fan and power supply handles, red for hot isle and blue for cold isle; a very nice and unique industrial design detail.

In Arista's product literature it states that typical power consumption of less than 2 Watt/port with twinax copper cables, and less than 3 Watt/port with SFP/QSFP lasers, the 7050S-64 provides industry leading power efficiency for the data center. We concur with this claim and its power consumption values.

© Lippis Enterprises, Inc. 2011

Discussion

The Arista 7050S-64 offers 64-wire speed 10GbE ports of layer 2 and 3 cut-through or store and forward mode forwarding in a 1 RU footprint. It can also be configured with 48 ports of 10GbE and 4-40GbE uplink ports. Designed with performance, low power and forward migration to 40GbE in mind, the Arista 7050S-64 provides low latency, line-rate forwarding and unique dynamic buffer allocation to keep traffic moving under congestion and micro-burst conditions. These claims were verified in the Lippis/Ixia RFC2544 latency test where average latency across all packet sizes was measured between 905 to 1,310 ns for layer 2 and 3 forwarding. Of all 64 10GbE port switches tested so far, the Arista 75050S-64 offers the lowest latency. Further, there was minimal delay variation, between 5 and 9 ns, across all packet sizes, demonstrating the switch architecture's ability to forward packets consistently and quickly boosting confidence that the Arista 7050S-64 will perform well in converged I/O configurations and consistently in demanding data center environments. In addition, the Arista 7050S-64 demonstrated 100% throughput at all packet sizes. Redundant power and fans along with various EOS high availability features ensure that the Arista 7050S-64 is always available and reliable for data center operations.

The 7050S-64 is a single chip 10GbE product, which provides consistent throughput and latency over all packet ranges as verified during the Lippis/Ixia test at iSimCity. The 7050S-64 is but one ToR switch in the Arista fixed configuration ToR products. In addition Arista also offers the 7100 ToR switches for ultra-low latency requirements. These Arista ToR switch products connect into Arista's 7500 Series of modular core switches, which provide industry-leading density of 192 to 384 10GbE connections. Across all Arista switches is a single binary image of its EOS operation system allowing all switches to utilize and employ its high availability features as well as unique operational modules including VM Tracer, vEOS and ZTP.

The combination of Arista's ToR and Core switches enable data center architects to build a flat, non blocking two-tier network with high 10GbE density, low latency, low power consumption and high availability. All Arista ToR and Core switches support 32 link Multi-Chassis LAG or Link Aggregation Group which enable high levels of switch interconnect to be built creating large scale out cloud computing facilities. The 7050S-64 is well suited as a ToR switch for high-density server connections where it would connect into Arista 7500 Core/Spine switches. This configuration scales up to 18,000 10GbE connections in a non-blocking two-tier architecture. Alternatively, for 1,500 to 3,000 10GbE connections the 7050S-64 can be both ToR/leaf and Core/Spine.

The 7050S-64 is well suited for low latency applications, storage interconnect, virtualized environment and other 10GbE data center deployments. Where high performance, low latency and low power consumption are all high priority attributes, the 7050S-64 should do very well.

© Lippis Enterprises, Inc. 2011

BLADE Network Technologies, an IBM Company — IBM BNT RackSwitch G8124

BLADE Network Technologies is an IBM company; referred here forward as BLADE. BLADE submitted two ToR switches for evaluation in the Lippis/Ixia test at iSimCity. Both products are based on a single chip 10GbE network design, which provides consistent throughput and latency over all packet ranges. Each product is profiled here along with their resulting test data.

The IBM BNT RackSwitch G8124 is a 24-port 10GbE switch, specifically designed for the data center, providing a virtual, cooler and easier network solution, according to BLADE. The IBM BNT RackSwitch G8124 is equipped with enhanced processing and memory to support large-scale aggregation layer 10GbE switching environments and applications, such as high frequency financial trading and IPTV.

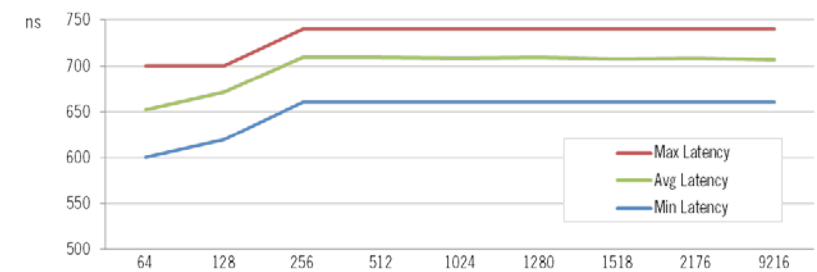
Video feature: Click to view IBM BNT RackSwitch G8124 video podcast

The IBM BNT RackSwitch G8124 ToR switch was tested across all 24 ports of 10GbE. Its average latency ranged from a low of 651 ns to a high of 709 ns for layer 2 traffic. Its average delay variation ranged between 5 and 10 ns, providing consistent latency across all packet sizes at full line rate.

BLADE Network Technologies, an IBM Company
IBM BNT RackSwitch G8124 Test Configuration

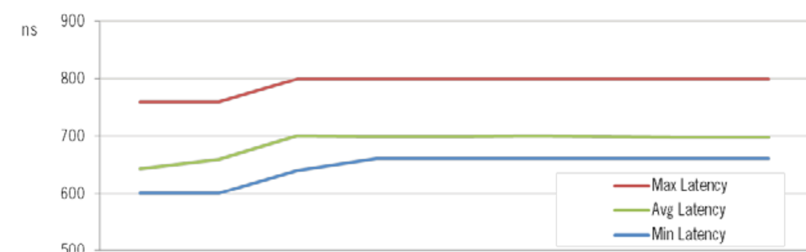
	Hardware	Software Version	Port Density
Device under test	IBM BNT RackSwitch G8124 http://www.bladenetwork.net/RackSwitch-G8124.html	6.5.2.3	24
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

IBM BNT RackSwitch G8124* RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	700	700	740	740	740	740	740	740	740
Avg Latency	651.75	671.5	709.458	709.042	708.25	709.708	707.917	708.042	706.625
Min Latency	600	620	660	660	660	660	660	660	660
Avg. delay Variation	9.25	5.083	5.417	7.875	7.167	5.25	10.125	5.167	10.125

IBM BNT RackSwitch G8124* RFC2544 Layer 3 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	760	760	800	800	800	800	800	800	800
Avg Latency	642.708	659.667	701.292	699.792	699	699.917	698.5	698.333	696.917
Min Latency	600	600	640	660	660	660	660	660	660
Avg. delay Variation	8.75	5.667	5.458	8.083	7.292	7.833	10.375	5.125	11.083

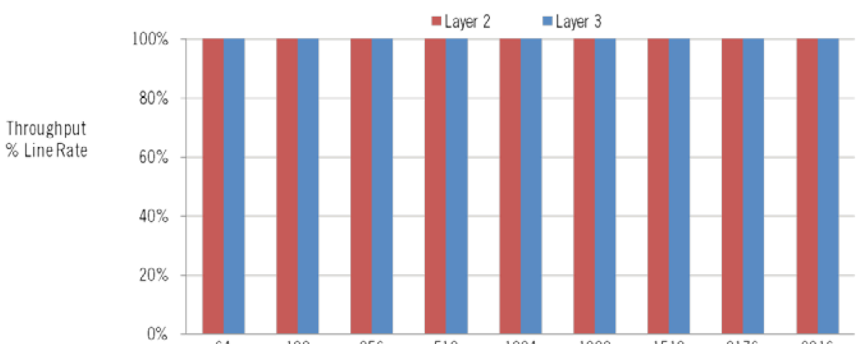
* The IBM BLADE switches were tested differently than all other switches. The IBM BNT RackSwitch G8264 and G8124 were configured and tested via cut-through test method, while all other switches were configured and tested via store-and-forward method. During store-and-forward testing, test equipment subtract packet transmission latency, decreasing actual latency measurements by the time it takes to transmit a packet from input to output port. Note that other potential device specific factors can impact latency too. This makes comparisons between two testing methodologies difficult.

© Lippis Enterprises, Inc. 2011

For 64Byte size packet, the Ixia test gear was configured for a staggered start.

For layer 3 traffic, the IBM BNT RackSwitch G8124's average latency ranged from a low of 642 ns to a high of 701 ns across all frame sizes. Its average delay variation for layer 3 traffic ranged between 5 ns and 11 ns, providing consistent latency across all packet sizes at full line rate.

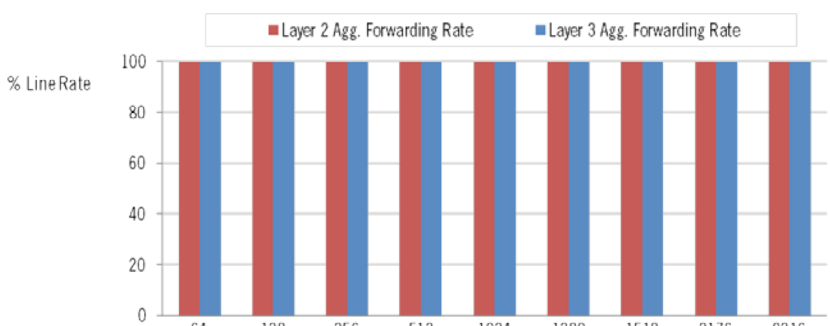
IBM BNT RackSwitch G8124 RFC 2544 L2 & L3 Throughput Test



	64	128	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100

The IBM BNT RackSwitch G8124 demonstrated 100% throughput as a percentage of line rate across all 24 10GbE ports. In other words, not a single packet was dropped while the IBM BNT RackSwitch G8124 was presented with enough traffic to populate its 24 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

IBM BNT RackSwitch G8124 RFC 2889 Congestion Test

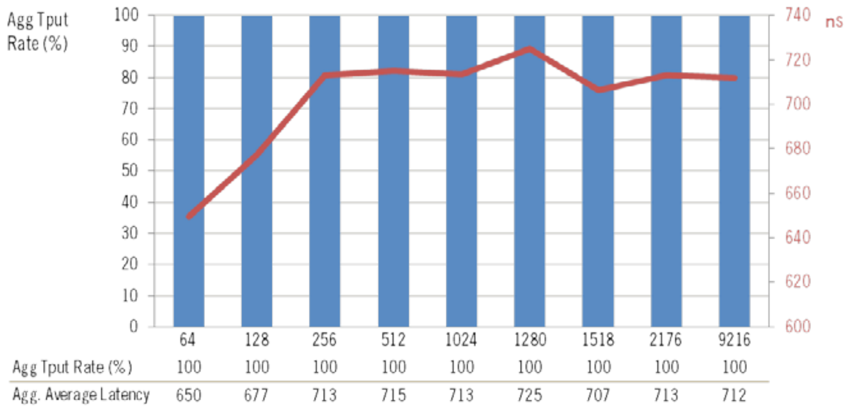


	64	128	256	512	1024	1280	1518	2176	9216
Layer 2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Layer 3 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Head of Line Blocking	no	no	no	no	no	no	no	no	no
Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes
L2 Agg Flow Control Frames	8202900	6538704	4531140	4303512	5108166	9139344	8690688	8958128	9022618
L3 Agg Flow Control Frames	8291004	6580984	4531070	4315590	4398264	4661968	4433112	4630880	4298994

The IBM BNT RackSwitch G8124 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions for both L2 and L3 traffic flows. A single 10GbE port was flooded at 150% of line rate. The IBM BNT RackSwitch G8124 did not exhibit Head of Line (HOL) blocking problems, which means that as the IBM BNT RackSwitch G8124-10GbE port became congested, it did not impact performance of other ports. As with most ToR switches, the IBM BNT RackSwitch G8124 did use back pressure as the Ixia test gear detected flow control frames.

© Lippis Enterprises, Inc. 2011

IBM BNT RackSwitch G8124 RFC 3918 IP Multicast Test



The IBM BNT RackSwitch G8124 demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a 650 ns at 64Byte size packet to 715 ns at 512 Byte size packets.

IBM BNT RackSwitch G8124 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	613
East-West	Server_to_Database	598
East-West	HTTP	643
East-West	iSCSI-Server_to_Storage	600
East-West	iSCSI-Storage_to_Server	619
North-South	Client_to_Server	635
North-South	Server_to_Client	581

The IBM BNT RackSwitch G8124 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 581 ns. Of special note is the consistency in latency across all traffic profiles observed during this test of the IBM BNT RackSwitch G8124.

IBM BNT RackSwitch G8124 Power Consumption Test

Watts_{ATIS}/10GbE port	5.5
3-Year Cost/Watts_{ATIS}/10GbE	$20.11
Total power cost/3-Year	$482.59
3 yr energy cost as a % of list price	4.04%
TEER Value	172
Cooling	Front to Back

The IBM BNT RackSwitch G8124 represents a new breed of cloud network leaf or ToR switches with power efficiency being a core value. Its Watts_{ATIS}/port is 5.5 and TEER value is 172. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. Its power cost per 10GbE is estimated at $6.70 per year. The three-year cost to power the IBM BNT RackSwitch G8124 is estimated at $482.59 and represents approximately 4% of its list price. Keeping with data center best practices, its cooling fans flow air front to back.

© Lippis Enterprises, Inc. 2011

Discussion:

The IBM BNT RackSwitch G8124 offers 24-wire speed 10GbE ports in a 1 RU footprint. Designed with performance in mind, the IBM BNT RackSwitch G8124 provides line-rate, high-bandwidth switching, filtering and traffic queuing without delaying data, and large data-center grade buffers to keep traffic moving. This claim was verified in the Lippis/Ixia RFC2544 latency test where average latency across all packet sizes was measured between 651 to 709 ns for layer 2 and 642 to 701 ns for layer 3 forwarding. Further, there was minimal delay variation, between 5 and 10 ns, across all packet sizes, demonstrating the switch architecture's ability to forward packets consistently and quickly. In addition, the IBM BNT RackSwitch G8124 demonstrated 100% throughput at all packet sizes independent upon layer 2 or layer 3 forwarding.

Redundant power and fans along with various high availability features ensure that the IBM BNT RackSwitch G8124 is always available for business-sensitive traffic. The low latency offered by the IBM BNT RackSwitch G8124 makes it ideal for latency sensitive applications, such as high performance computing clusters and financial applications. Further, the IBM BNT RackSwitch G8124 supports the newest data center protocols including CEE for support of FCoE which was not evaluated in the Lippis/Ixia test.

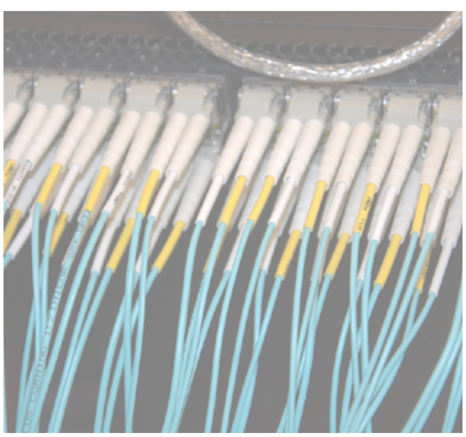
© Lippis Enterprises, Inc. 2011

BLADE Network Technologies, an IBM Company — IBM BNT RackSwitch G8264

The IBM BNT RackSwitch G8264 is a wire speed 10 and 40GbE switch specifically designed for the data center. The IBM BNT RackSwitch G8264 is offered in two configurations: either as a 64-port wire speed 10GbE ToR switch or 48-port wire speed 10GbE with four-wire speed 40GbE ports. The total bandwidth of the IBM BNT RackSwitch G8264 is 1.2 Terabits per second packaged in a 1 RU footprint. For the Lippis/Ixia test, the IBM BNT RackSwitch G8264 was configured as a 64-port wire speed 10GbE ToR switch.

The IBM BNT RackSwitch G8264 ToR switch was tested across all 64 ports of 10GbE. Its average latency ranged from a low of 1081 ns to a high of 1421 ns for layer 2 traffic. Its average delay variation ranged between 5 and 9.5 ns, providing consistent latency across all packet sizes at full line rate.

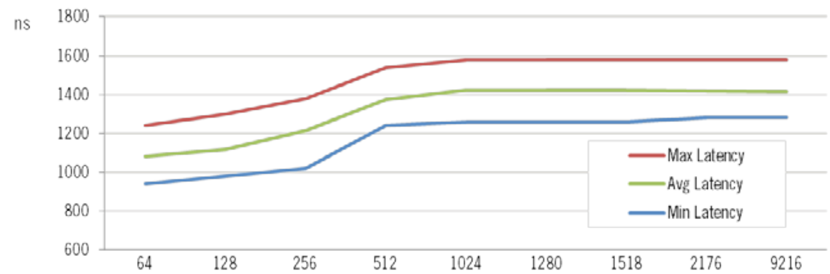
For layer 3 traffic, the IBM BNT RackSwitch G8264's average latency ranged from a low of 1075 ns to a high of 1422 ns across all frame sizes. Its average delay variation for layer 3 traffic ranged between 5 and 9.5 ns, providing consistent latency across all packet sizes at full line rate.



BLADE Network Technologies, an IBM Company
IBM BNT G8264 Test Configuration

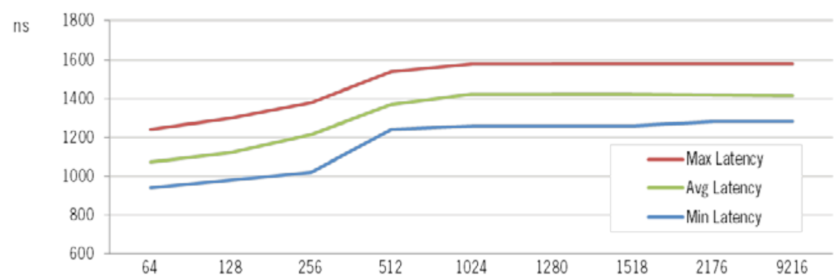
	Hardware	Software Version	Port Density
Device under test	IBM BNT RackSwitch G8264 http://www.bladenetwork.net/RackSwitch-G8264.html	6.4.1.0	64
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

IBM BNT RackSwitch G8264*
RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	1240	1300	1380	1540	1580	1580	1580	1580	1580
Avg Latency	1081.344	1116.828	1214.922	1374.25	1422	1421.281	1421.75	1419.891	1415.359
Min Latency	940	980	1020	1240	1260	1260	1260	1280	1280
Avg. delay Variation	8.688	5.219	5.609	7.609	6.922	7.563	9.641	5.047	9.563

IBM BNT RackSwitch G8264*
RFC2544 Layer 3 Latency Test

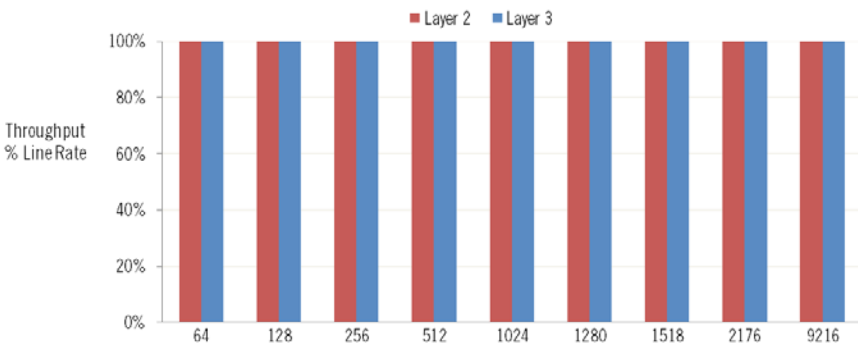


	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	1240	1300	1380	1540	1580	1580	1580	1580	1580
Avg Latency	1075.391	1121.016	1213.547	1370.938	1421.938	1422.625	1421.828	1419.734	1415.781
Min Latency	940	980	1020	1240	1260	1260	1260	1280	1280
Avg. delay Variation	8.609	5.172	5.641	7.672	6.734	6.25	9.422	5.078	9.578

* The IBM BLADE switches were tested differently than all other switches. The IBM BNT RackSwitch G8264 and G8124 were configured and tested via cut-through test method, while all other switches were configured and tested via store-and-forward method. During store-and-forward testing, test equipment subtract packet transmission latency, decreasing actual latency measurements by the time it takes to transmit a packet from input to output port. Note that other potential device specific factors can impact latency too. This makes comparisons between two testing methodologies difficult.

© Lippis Enterprises, Inc. 2011

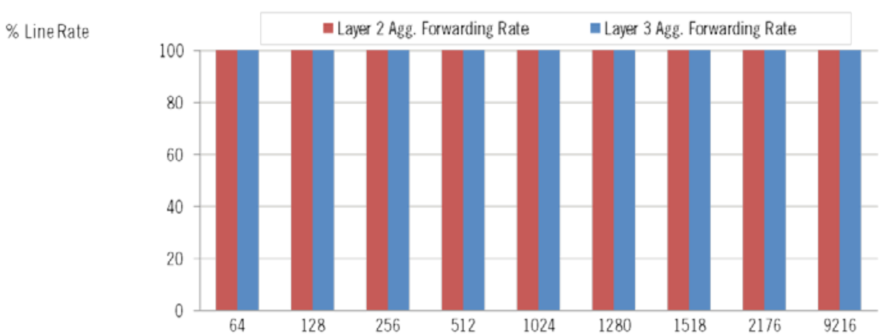
IBM BNT RackSwitch G8264
RFC 2544 L2 & L3 Throughput Test



	64	128	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100

The IBM BNT RackSwitch G8264 demonstrated 100% throughput as a percentage of line rate across all 64-10GbE ports. In other words, not a single packet was dropped while the IBM BNT RackSwitch G8264 was presented with enough traffic to populate its 64 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

IBM BNT RackSwitch G8264
RFC 2889 Congestion Test

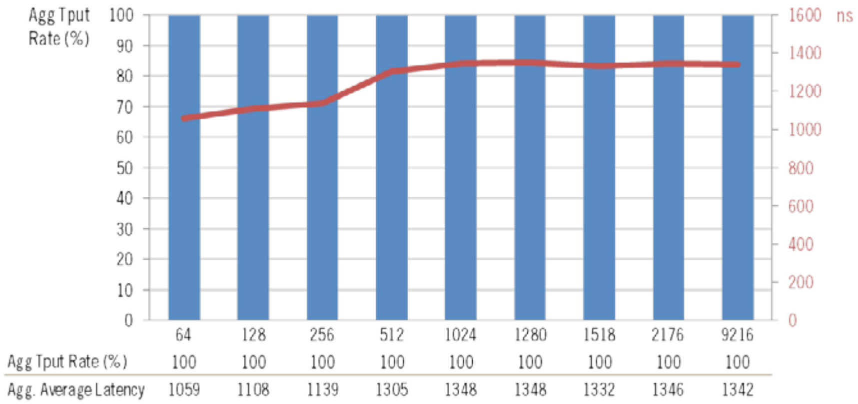


	64	128	256	512	1024	1280	1518	2176	9216
Layer 2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Layer 3 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Head of Line Blocking	no	no	no	no	no	no	no	no	no
Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes
L2 Agg Flow Control Frames	7498456	5582384	6014492	5352004	3751186	514182	844520	903536	9897178
L3 Agg Flow Control Frames	7486426	5583128	6008396	5351836	5348868	5096720	4982224	4674228	3469558

The IBM BNT RackSwitch G8264 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions for both L2 and L3 traffic flows. A single 10GbE port was flooded at 150% of line rate. The IBM BNT RackSwitch G8124 did not exhibit Head of Line (HOL) blocking problems, which means that as the IBM BNT G8264 10GbE port became congested, it did not impact the performance of other ports. As with most ToR switches, the IBM BNT RackSwitch G8264 did use back pressure as the Ixia test gear detected flow control frames.

© Lippis Enterprises, Inc. 2011

IBM BNT RackSwitch G8264 RFC 3918 IP Multicast Test



The IBM BNT RackSwitch G8264 demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a 1059 ns at 64Byte size packet to 1348 ns at 1024 and 1280 Byte size packets.

IBM BNT RackSwitch G8264 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	1060
East-West	Server_to_Database	1026
East-West	HTTP	1063
East-West	iSCSI-Server_to_Storage	1026
East-West	iSCSI-Storage_to_Server	1056
North-South	Client_to_Server	1058
North-South	Server_to_Client	1017

The IBM BNT RackSwitch G8264 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 1017 ns. Of special note is the consistency in latency across all traffic profiles observed during this test of the IBM BNT RackSwitch G8264.

IBM BNT RackSwitch G8264 Power Consumption Test

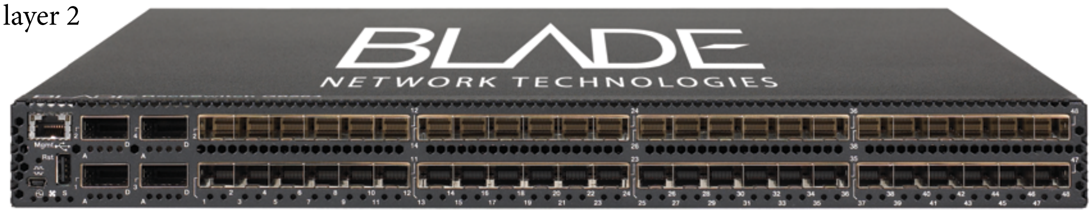
$Watts_{ATIS}$/10GbE port	3.92
3-Year Cost/$Watts_{ATIS}$/10GbE	$14.33
Total power cost/3-Year	$917.22
3 yr energy cost as a % of list price	4.08%
TEER Value	241
Cooling	Front to Back

The IBM BNT RackSwitch G8264 represents a new breed of cloud network leaf or ToR switches with power efficiency being a core value. Its $Watts_{ATIS}$/port is 3.92 and TEER value is 241. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. Its power cost per 10GbE is estimated at $4.78 per year. The three-year cost to power the IBM BNT RackSwitch G8264 is estimated at $917.22 and represents approximately 4% of its list price. Keeping with data center best practices, its cooling fans flow air front to back.

© Lippis Enterprises, Inc. 2011

Discussion:

The IBM BNT RackSwitch G8264 was designed with top performance in mind as it provides line-rate, high-bandwidth switching, filtering, and traffic queuing with low consistent latency. This claim was verified in the Lippis/Ixia RFC2544 latency test, where average latency across all packet sizes was measured between 1081 to 1422 ns for layer 2 and 1075 to 1422 ns for layer 3 forwarding. Further, there was minimal delay variation, between 5 and 9 ns, across all packet sizes for layer 2 and 3 forwarding, demonstrating the switch's architecture to forward packets at nanosecond speeds consistently under load. In addition, the IBM BNT RackSwitch G8264 demonstrated 100% throughput at all packet sizes independent upon layer 2 or layer 3 forwarding.



The low latency offered by the IBM BNT RackSwitch G8264 makes it ideal for latency sensitive applications, such as high performance computing clusters and financial applications. In fact, the IBM BNT RackSwitch G8264 demonstrated 100% aggregated throughput at all packet sizes with latency measured in the 1059 to 1348 ns range during the Lippis/Ixia RFC3918 IP multicast test.

Nanosecond latency offered by the IBM BNT RackSwitch G8264 makes it ideal for latency sensitive applications, such as high performance computing clusters and financial applications. Further, the IBM BNT RackSwitch G8264 supports the newest data center protocols including CEE for support of FCoE which was not evaluated in the Lippis/Ixia test.

© Lippis Enterprises, Inc. 2011

Brocade VDX™ 6720-24

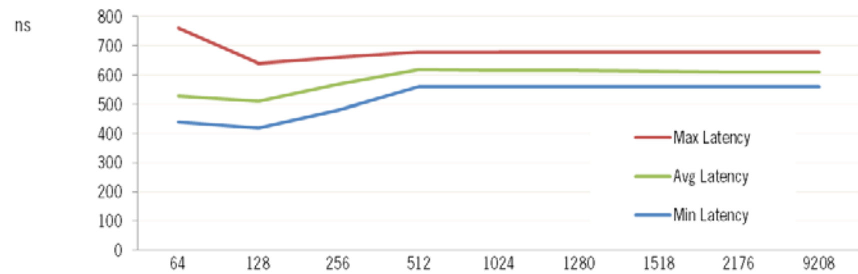
For this Lippis/Ixia evaluation at iSimCity Brocade submitted the industry's first fabric enabled VDX™ 6720-24 Top-of-Rack (ToR) switch that's built with Brocade's Virtual Cluster Switching (VCS™) technology, an Ethernet fabric innovation that addresses the unique requirements of evolving data center environments. The Brocade VDX™ 6720-24 is a 24 port 10 Gigabit Ethernet (10GbE) ToR switch. Brocade® VDX™ 6720 Data Center Switches, are 10 GbE line-rate, low-latency, lossless switches in 24-port (1U) and 60-port (2U) models. The Brocade VDX™ 6720 series switch is based on sixth-generation ASIC fabric switching technology and runs the feature dense Brocade Network Operating System (NOS).

The VDX™ 6720-24 ToR switch was tested across all 24 ports of 10GbE. Its average latency ranged from a low of 510 ns to a high of 619 ns for layer 2 traffic. Its average delay variation ranged between 5 and 8 ns, providing consistent latency across all packet sizes at full line rate.

Brocade VDX™ 6720-24 Test Configuration

	Hardware	Software Version	Port Density
Device under test	VDX™ 6720-24 http://www.brocade.com/	2.0.0a	24
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Brocade VDX™ 6720-24 RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9208
Max Latency	760	640	660	680	680	680	680	680	680
Avg Latency	527.042	510.708	570.083	619.458	615.542	615.583	614.208	612.167	610.958
Min Latency	440	420	480	560	560	560	560	560	560
Avg. delay Variation	8.75	5.375	5.875	7.75	6.375	8.25	8.75	5.167	6.542

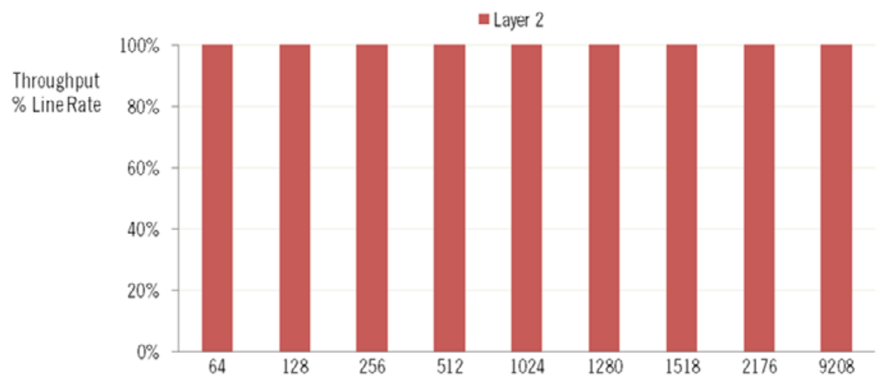
note max frame size is 9208 so could not do 9216

Brocade VDX™ 6720-24 RFC2544 Layer 3 Latency Test

The Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test.

© Lippis Enterprises, Inc. 2011

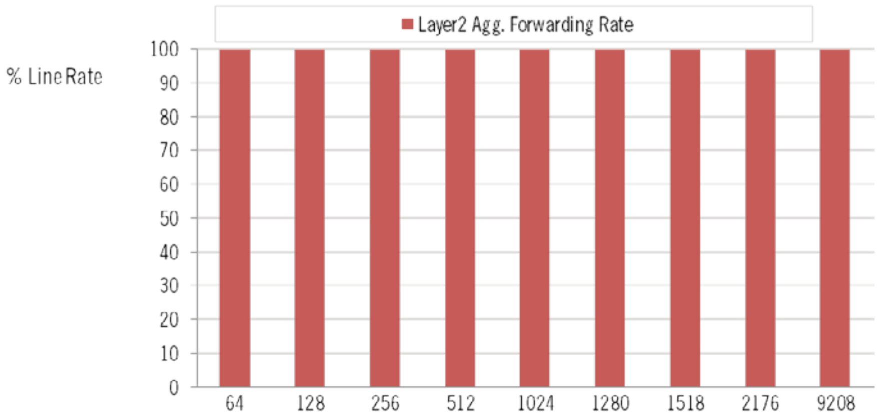
Brocade VDX™ 6720-24
RFC 2544 L2 Throughput Test



	64	128	256	512	1024	1280	1518	2176	9208
Layer 2	100	100	100	100	100	100	100	100	100

The Brocade VDX™ 6720-24 demonstrated 100% throughput as a percentage of line rate across all 24 10GbE ports. In other words, not a single packet was dropped while the Brocade VDX™ 6720-24 was presented with enough traffic to populate its 24 10GbE ports at line rate simultaneously for L2 traffic flows.

Brocade VDX™ 6720-24
RFC 2889 Congestion Test



	64	128	256	512	1024	1280	1518	2176	9208
Layer2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
L2 Head of Line Blocking	no	no	no	no	no	no	no	no	no
L2 Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes

The Brocade VDX™ 6720-24 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions. A single 10GbE port was flooded at 150% of line rate. The Brocade VDX™ 6720-24 did not exhibit Head of Line (HOL) blocking problems, which means that as the Brocade VDX™ 6720-24 port became congested, it did not impact performance of other ports. As with most ToR switches, the Brocade VDX™ 6720-24 uses back pressure as the Ixia test gear detected flow control frames.

© Lippis Enterprises, Inc. 2011

Brocade VDX™ 6720-24 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	7653
East-West	Server_to_Database	675
East-West	HTTP	4805
East-West	iSCSI-Server_to_Storage	1298
East-West	iSCSI-Storage_to_Server	7901
North-South	Client_to_Server	1666
North-South	Server_to_Client	572

The Brocade VDX™ 6720-24 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 7901 ns, measured in cut-through mode. Of special note is the low latency of 572 ns and 675 ns observed during north-south server-to-client and east-west server-to-database flows respectively.

Brocade VDX™ 6720-24 Power Consumption Test

$Watts_{ATIS}$/10GbE port	3.05
3-Year Cost/$Watts_{ATIS}$/10GbE	$11.15
Total power cost/3-Year	$267.62
3 yr energy cost as a % of list price	2.24%
TEER Value	310
Cooling	Front to Back

The Brocade VDX™ 6720-24 represents a new class of Ethernet fabric enabled fixed port switches with power efficiency being a core value. Its WattsATIS/port is a very low 3.05 and TEER value is 310. Note TEER values represent the ratio of work performed over energy consumption. Hence, higher TEER values are more desirable. Its power cost per 10GbE is estimated at $3.72 per year. The three-year cost to power the Brocade VDX™ 6720-24 is estimated at $267.62 and represents approximately 2.24% of its list price. Keeping with data center best practices, its cooling fans flow air rear to front. There is also a front to back airflow option available to meet all customer cooling requirements.

© Lippis Enterprises, Inc. 2011

Discussion:

New Ethernet fabric based architectures are emerging to support highly virtualized data center environments and the building of cloud infrastructures allowing IT organizations to be more agile and customer responsive. Ethernet fabrics will transition the data center networking from the inefficient and cumbersome Spanning Tree Protocol (STP) networks, to high performance, flat, layer 2 networks built for the east-west traffic demanded by server virtualization and highly clustered applications.

The Brocade VDX™ 6720-24 with VCS technology offers 24-wire speed 10GbE ports of layer 2 cut-through forwarding in a 1 RU footprint. Designed with performance in mind, the Brocade VDX™ 6720-24 provides line-rate, high-bandwidth switching, filtering and traffic queuing without delaying data, and large data-center grade buffers to keep traffic moving. This claim was verified in the Lippis/Ixia RFC2544 latency test where average latency across all packet sizes was measured between 510 to 619 ns for layer 2 forwarding. Further, there was minimal delay variation, between 5 and 8 ns, across all packet sizes, demonstrating the switch architecture's ability to forward packets consistently and quickly boosting confidence that the Brocade VDX™ 6720-24 will perform well in converged I/O configurations. In addition, the Brocade VDX™ 6720-24 demonstrated 100% throughput at all packet sizes. Redundant power and fans along with various high availability features ensure that the Brocade VDX™ 6720-24 is always available and reliable for data center operations.

The VDX™ 6720-24 is build on innovative Brocade designed ASIC, which provides consistent throughput and latency over all packet ranges as verified during the Lippis/Ixia test at iSimCity. The VDX™ 6720-24 is the entry level of the VDX family, which can be configured as a single logical chassis and scales to support 16-, 24-, 40-, 50-, and 60-ports of 10 GbE.

Brocade VCS capabilities allow customers to start from small two-switch ToR deployments, such as the VDX™ 6720-24 tested here, and scale to large virtualization-optimized Ethernet fabrics with tens of thousands of ports overtime. Furthermore, Brocade claims that VCS-based Ethernet fabrics are convergence ready, allowing customers to deploy one network for storage—Fibre Channel over Ethernet (FCoE), iSCSI, NAS—and IP data traffic.

Brocade VCS technology boasts lower capital cost by collapsing access and aggregation layers of a three-tier network into two and thus reducing operating cost through managing fewer devices. Customers will be able to seamlessly transition their existing multi-tier hierarchical networks to a much simpler, virtualized, and converged data center network—moving at their own pace, according to Brocade.

© Lippis Enterprises, Inc. 2011

Force10 S-Series S4810

The Force10 S-Series S4810 is a low-latency (nanosecond scale) 10/40 GbE ToR/leaf switch purpose-built for applications in high-performance data center and computing environments. The compact S4810 design provides port density of 48 dual-speed 1/10 GbE (SFP+) ports as well as four 40 GbE QSFP+ uplinks to conserve valuable rack space and simplify the migration to 40 Gbps in the data center core. For the Lippis/Ixia test, the S4810 was configured as a 48-wire speed 10 GbE ToR switch.

Video feature: Click to view Force10 video podcast

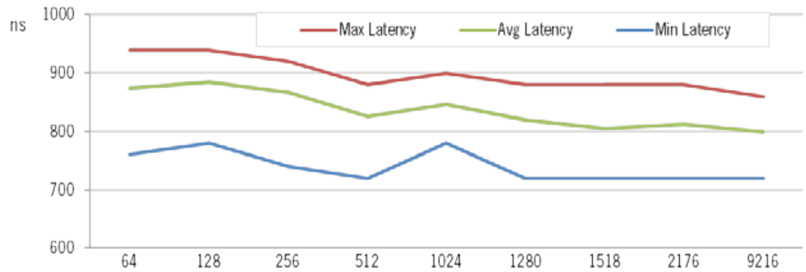
The Force10 S4810 ToR switch was tested across all 48 ports of 10GbE. Its average latency ranged from a low of 799 ns to a high of 885 ns for layer 2 traffic. Its average delay variation ranged between 5 and 9.2 ns, providing consistent latency across all packet sizes at full line rate.

For layer 3 traffic, the Force10 S4810's average latency ranged from a low of 796 ns to a high of 868 ns across all frame sizes. Its average delay variation for layer 3 traffic ranged between 4.6 and 9.1 ns, providing consistent latency across all packet sizes at full line rate.

Force10 Networks S4810 Test Configuration

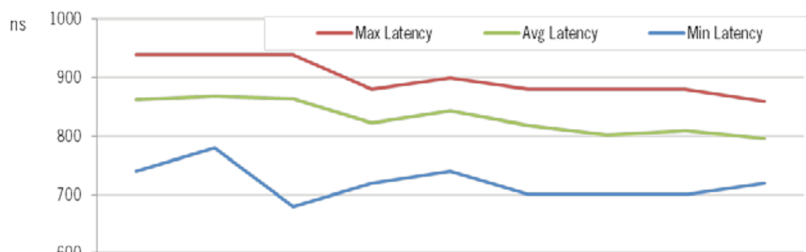
	Hardware	Software Version	Port Density
Device under test	S4810 http://www.force10networks.com	8.3.7.0E4	48
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Force10 Networks S4810 RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	940	940	920	880	900	880	880	880	860
Avg Latency	874.479	885	867.604	825.292	846.521	819.896	805.063	812.063	799.792
Min Latency	760	780	740	720	780	720	720	720	720
Avg. delay Variation	9.104	5.229	5.458	7.938	6.729	6	8.979	5	9.271

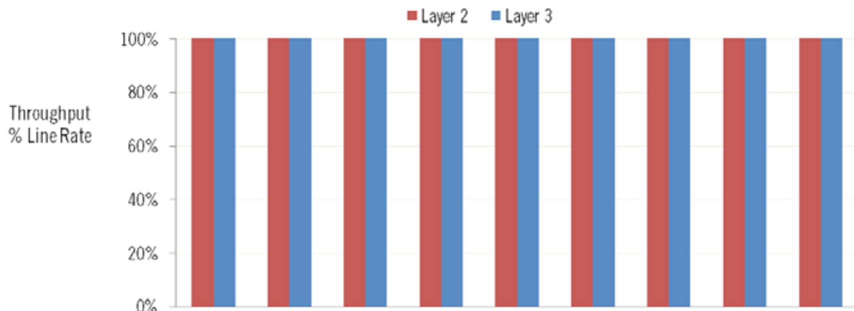
Force10 Networks S4810 RFC2544 Layer 3 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	940	940	940	880	900	880	880	880	860
Avg Latency	862.417	868.542	863.771	822.625	843.021	817.833	802.375	809.708	796.063
Min Latency	740	780	680	720	740	700	700	700	720
Avg. delay Variation	8.458	5.25	5.521	7.688	6.688	4.646	9.063	5.083	9.146

© Lippis Enterprises, Inc. 2011

Force10 Networks S4810
RFC 2544 L2 & L3 Throughput Test



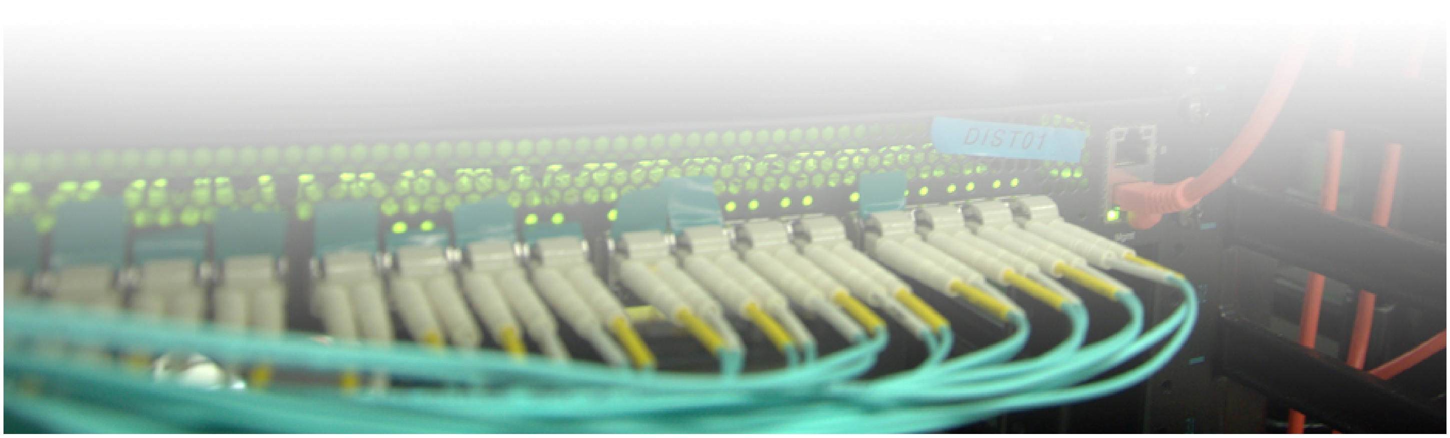
	64	128	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100

The Force10 S4810 demonstrated 100% throughput as a percentage of line rate across all 48-10GbE ports. In other words, not a single packet was dropped while the Force10 S4810 was presented with enough traffic to populate its 48 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

Force10 Networks S4810
RFC 2889 Congestion Test

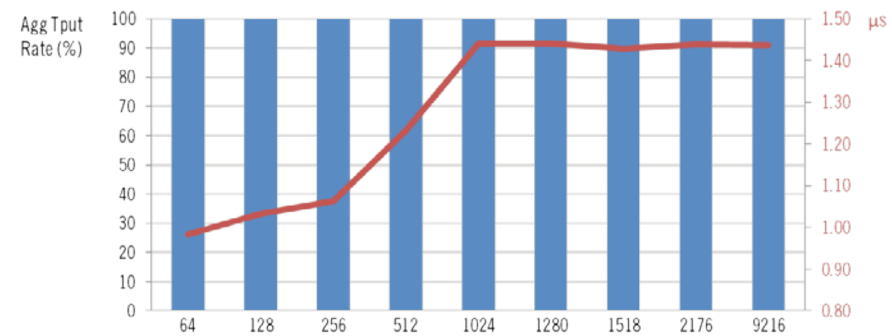
	64	128	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Layer 3 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Head of Line Blocking	no	no	no	no	no	no	no	no	no
Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes
L2 Agg Flow Control Frames	5761872	3612428	3841000	3073936	3158216	2958772	2876596	2730684	2681572
L3 Agg Flow Control Frames	5762860	3614948	3851964	3070936	3155552	2957128	2875592	2730528	2677416

The Force10 S4810 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions for both L2 and L3 traffic flows. A single 10GbE port was flooded at 150% of line rate. The Force10 S4810 did not use HOL blocking, which means that as the Force10 S4810-10GbE port became congested, it did not impact the performance of other ports. As with most ToR switches, the Force10 S4810 did use back pressure as the Ixia test gear detected flow control frames.



© Lippis Enterprises, Inc. 2011

Force10 Networks S4810
RFC 3918 IP Multicast Test



	64	128	256	512	1024	1280	1518	2176	9216
Agg Tput Rate (%)	100	100	100	100	100	100	100	100	100
Agg. Average Latency	0.984	1.033	1.062	1.232	1.442	1.442	1.429	1.438	1.437

The Force10 S4810 demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a 984 ns at 64Byte size packet to 1442 ns at 1024 and 1280 Byte size packets. The Force10 S4810 was configured in cut-through mode during this test.

Force10 Networks S4810 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	6890
East-West	Server_to_Database	1010
East-West	HTTP	4027
East-West	iSCSI-Server_to_Storage	771
East-West	iSCSI-Storage_to_Server	6435
North-South	Client_to_Server	1801
North-South	Server_to_Client	334

The Force10 S4810 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 6900 ns. The Force10 S4810 was configured in cut-through mode during this test.

Force10 Networks S4810 Power Consumption Test

$Watts_{ATIS}$/10GbE port	4.03
3-Year Cost/$Watts_{ATIS}$/10GbE	$14.73
Total power cost/3-Year	$707.22
3 yr energy cost as a % of list price	2.83%
TEER Value	235
Cooling	Front to Back

The Force10 S4810 represents a new breed of cloud network leaf or ToR switches with power efficiency being a core value. Its $Watts_{ATIS}$/port is 4.03 and TEER value is 235. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. Its power cost per 10GbE is estimated at $4.91 per year. The three-year cost to power the Force10 S4810 is estimated at $707.22 and represents approximately 3% of its list price. Keeping with data center best practices, its cooling fans flow air front to back.

© Lippis Enterprises, Inc. 2011

Discussion:

Leveraging a non-blocking, cut-through switching architecture, the Force10 Networks S4810 delivers line-rate L2 and L3 forwarding capacity with nanosecond latency to maximize network performance. These claims were substantiated during the Lippis/Ixia RFC 2544 latency test, where the S4810 demonstrated average latency of 799 to 874 ns across all packet sizes for layer 2 forwarding and 796 to 862 ns across all packet sizes for layer 3 forwarding. The S4810's architecture demonstrated its ability to forward packets in the nanosecond range consistently with little delay variation. In fact, the S4810 demonstrated between 5 and 9 ns of delay variation across all packet sizes for both layer 2 and layer 3 forwarding. The S4810 demonstrated 100% throughput at line rate for all packet sizes ranging from 64 Bytes to 9216 Bytes.

In addition to the above verified and tested attributes, Force10 Networks claims that the S4810 is designed with powerful QoS features, coupled with Data Center Bridging (DCB) support via a future software enhancement, making the S4810 ideally suited for iSCSI storage environments. In addition, the S4810 incorporates multiple architectural features that optimize data center network flexibility, efficiency and availability, including Force10's VirtualScale stacking technology, reversible front-to-back or back-to-front airflow for hot/cold aisle environments, and redundant, hot-swappable power supplies and fans.

The S4810 also supports Force10's Open Automation Framework, which provides advanced network automation and virtualization capabilities for virtual data center environments. The Open Automation Framework is comprised of a suite of inter-related network management tools that can be used together or independently to provide a network that is flexible, available and manageable while reducing operational expenses.

© Lippis Enterprises, Inc. 2011

Hitachi Cable Apresia 15000-64XL-PSR

The Apresia 15000 series was developed under the new BoxCore concept for improving network efficiency by using box switches as Core switches to fulfill customers' satisfaction. The Apresia 15000-64XL-PSR, referred to here as the 15K, is a 64-port wire speed 10GbE Core switch. As the Apresia 15K is a 64-port 10GbE switch, it could be deployed in ToR, End-of-Row or as a Core/Spine switch. For the Lippis/Ixia evaluation, Hitachi Cable requested that the Apresia 15K be grouped with Core switches. For the Lippis/Ixia evaluation, Hitachi Cable requested that the Apresia 15K be grouped with Core switches, but is more appropriate for the ToR category.

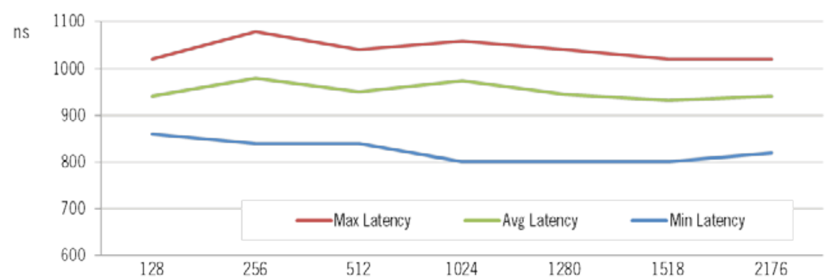
Video feature: Click to view Apresia video podcast

The Apresia 15K switch was tested across all 64 ports of 10GbE. Note that during testing, the Apresia 15K did not maintain VLAN through the switch at 64Byte size packets, which is used to generate a packet signature to measure latency. Therefore, there is no 64Byte latency test data available. Hitachi claims they have fixed the issue after the test. Further, the Apresia 15K's largest frame size supported in 9044 Bytes, therefore, precluding 9216 Byte size testing.

Apresia 15000-64XL-PSR Test Configuration

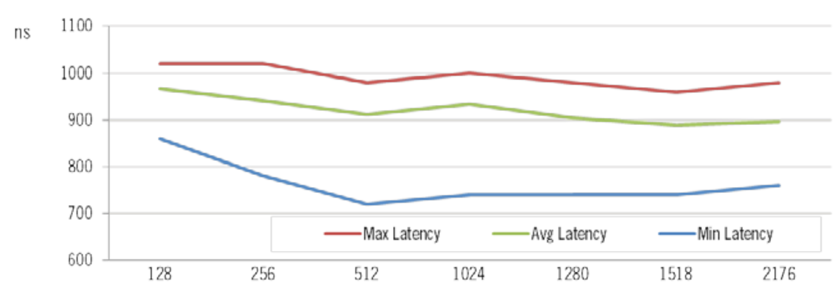
	Hardware	Software Version	Port Density
Device under test	15000-64XL-PSR http://www.apresia.jp/en/	AEOS 8.09.01	64
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Apresia 15000-64XL-PSR RFC2544 Layer 2 Latency Test



	128	256	512	1024	1280	1518	2176
Max Latency	1020	1080	1040	1060	1040	1020	1020
Avg Latency	940.672	979.953	950.063	973.547	944.641	932.516	940.344
Min Latency	860	840	840	800	800	800	820
Avg Latency Delay	7.984	7.063	8	9.453	7	10.031	8.031

Apresia 15000-64XL-PSR RFC2544 Layer 3 Latency Test



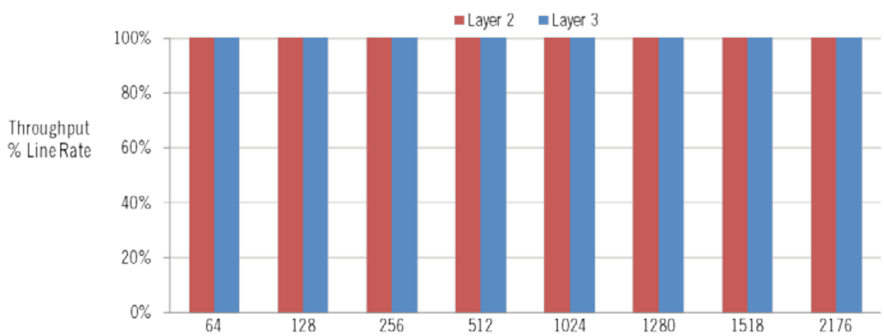
	128	256	512	1024	1280	1518	2176
Max Latency	1020	1020	980	1000	980	960	980
Avg Latency	967.172	941.078	910.844	933.047	904.844	889.547	897.328
Min Latency	860	780	720	740	740	740	760
Avg delay Variation	5.234	5.609	7.719	6.641	6.047	9.234	5.016

© Lippis Enterprises, Inc. 2011

Latency data is measured for packet sizes between 128 and 9044. Within this packet range, the Apresia 15K's average latency was a low of 900 ns to a high of 979 ns for layer 2 traffic. Its average delay variation ranged between 7 and 10.5 ns, providing consistent latency across these packet sizes at full line rate.

For layer 3 traffic, the Apresia 15K's average latency ranged from a low of 889 ns to a high of 967 ns across 128-9044 Bytes size frames. Its average delay variation for layer 3 traffic ranged between 5 and 9.6 ns, providing consistent latency across these packet sizes at full line rate.

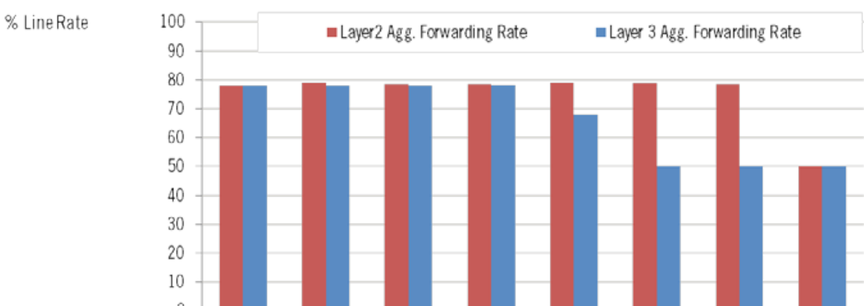
Apresia 15000-64XL-PSR
RFC 2544 L2 & L3 Throughput Test



	64	128	256	512	1024	1280	1518	2176
Layer 2	100	97.288	98.541	99.238	99.607	99.682	99.73	99.808
Layer 3	100	100	100	100	100	100	100	100

The Apresia 15K demonstrated nearly 100% throughput as a percentage of line rate across all 64-10GbE ports for layer 2 traffic and 100% throughput for layer 3 flows. In other words, there was some packet loss during the RFC 2544 Layer 2 test, but for Layer 2 not a single packet was dropped while the Apresia 15K was presented with enough traffic to populate its 64 10GbE ports at line rate simultaneously.

Apresia 15000-64XL-PSR
RFC 2889 Congestion Test



	64	128	256	512	1024	1280	1518	2176
Layer2 Agg. Forwarding Rate	77.854	79.095	78.624	78.581	79.029	78.974	78.609	49.946
Layer 3 Agg. Forwarding Rate	77.913	77.911	77.785	78.229	67.881	50.071	50.029	50.03
L2 Head of Line Blocking	no	no	no	no	no	no	no	yes
L3 Head of Line Blocking	no	no	no	no	yes	yes	yes	yes
L2 Back Pressure	yes	yes	yes	yes	yes	yes	yes	no
L3 Back Pressure	yes	yes	yes	yes	no	no	no	no
Agg Flow Control Frames	0	0	0	0	0	0	0	0

The Apresia 15K demonstrated mixed results for the congestion test. Between 64 and 1518 Byte size packets, the 15K demonstrated nearly 80% of aggregated forwarding rate as percentage of line rate during congestion conditions for L2 traffic flows. At 2176 Byte size packets, the aggregated forwarding rate dropped to nearly 50% with HOL blocking present.

For L3 traffic flows at 64 and 1512 Byte size packets, the 15K demonstrated nearly 80% of aggregated forwarding rate as percentage of line rate during congestion conditions. At packet sizes 1024 to 2176, HOL blocking was present as the aggregated forwarding rate as percentage of line rate during congestion conditions dropped to 67% and 50%.

© Lippis Enterprises, Inc. 2011

Apresia 15000-64XL-PSR
RFC 3918 IP Multicast Test

The Apresia 15K does not support IP Multicast.

Apresia 15K Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	1698
East-West	Server_to_Database	1200
East-West	HTTP	1573
East-West	iSCSI-Server_to_Storage	1148
East-West	iSCSI-Storage_to_Server	1634
North-South	Client_to_Server	2183
North-South	Server_to_Client	1021

The Apresia 15K delivered 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows during the cloud computing simulation test. Zero packet loss was observed as its latency stayed under 2183 ns.

Apresia 15K Power Consumption Test

$Watts_{ATIS}$/10GbE port	3.58
3-Year Cost/$Watts_{ATIS}$/10GbE	$13.09
Total power cost/3-Year	$837.67
TEER Value	264
Cooling	Front to Back

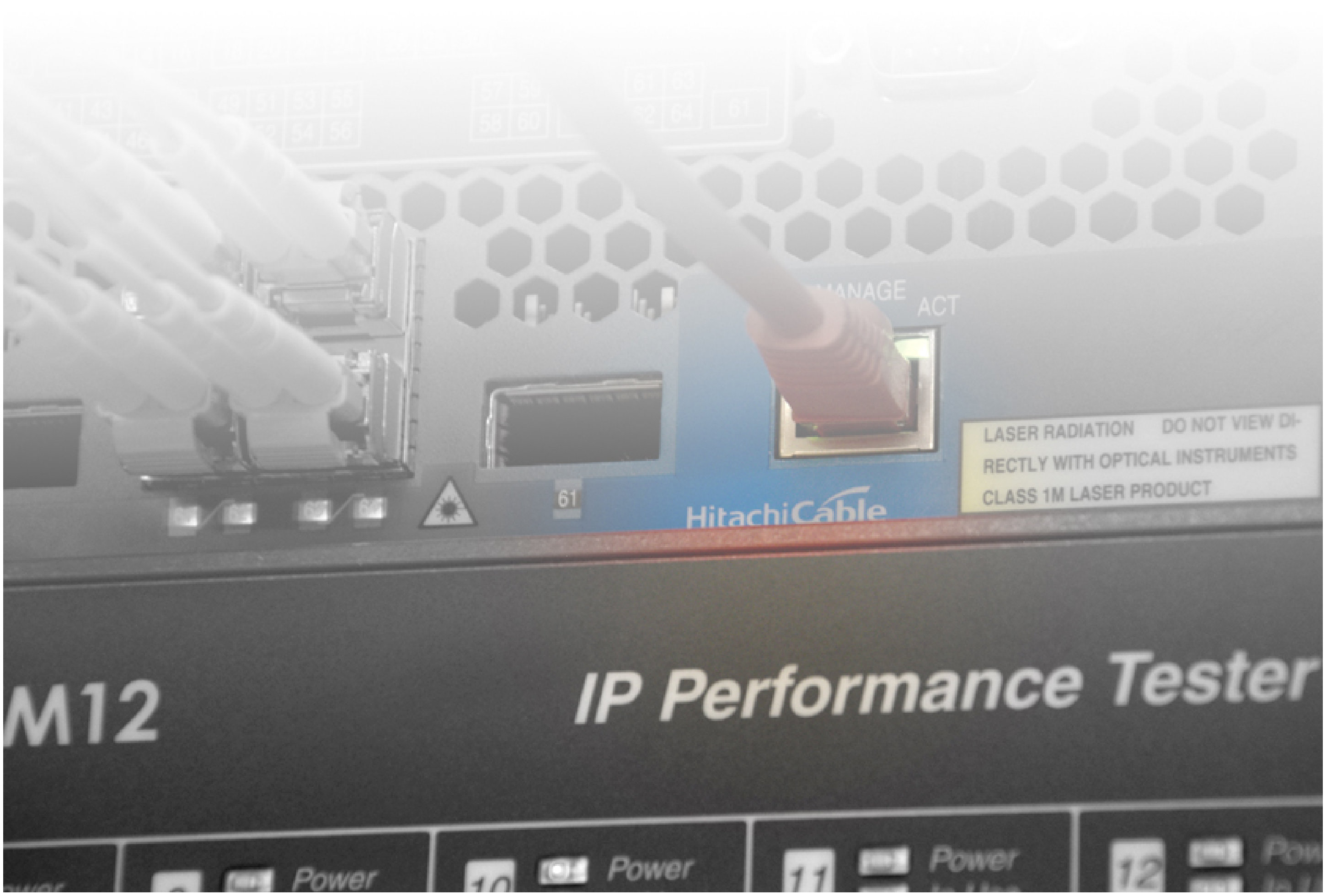
The Apresia 15K is a new breed of cloud network ToR or leaf switch with power efficiency being a core value. Its $Watts_{ATIS}$/port is 3.58 and TEER value is 264. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. Hitachi Cable was unable to provide list pricing information thus there is no data on the Apresia 15K's power cost per 10GbE, three-year cost of power and estimate of three-year energy cost as a percentage of list price. Keeping with data center best practices, its cooling fans flow air front to back and states that Back-to-Front cooling will be available at some point in the future.

© Lippis Enterprises, Inc. 2011

Discussion:
The Apresia 15000-64XL-PSR, referred to here as the 15K

The Apresia 15000-64XL-PSR (hereinafter called 64XL) is a terabit box switch that implements 2 40G uplink ports, and 64 1G/10G SFP/SFP+ ports in a 2U size unit. The Apresia15000-32XL-PSR (hereinafter called 32XL) is half the size of 64XL and implements 2 40G uplink ports, and 32 1G/10G SFP/SFP+ ports in a 1U size unit. Switching capacities are 1.28 terabits/second for the 64XL, and 640 gigabits/second for the 32XL.

When set to equal capacities, the 64XL and the 32XL are not only space-saving, but also achieve significant cost-savings when compared to the conventional chassis switches. Both the 64XL and 32XL are designed to be used as data center switches, or network Core switches and broadband L2 switches for enterprises and academic institutions. Planned functions specific to data center switches that play an important role in cloud computing include FCoE, storage I/O consolidation technology, and DCB, a new Ethernet technical standard to realize FCoE, and other functions that support virtual server environments. A data center license and L3 license are optional and can be purchased in accordance with the intended use. Without such licenses, both the 64XL and 32XL perform as high-end L2 switches.



© Lippis Enterprises, Inc. 2011

Juniper Network EX Series EX8200 Ethernet Switch, EX8216 Ethernet Switch

The 16-slot EX8216 Ethernet Switch, part of the EX8200 line of Ethernet Switches from Juniper Networks®, offers a high-density, high-performance platform for aggregating access switches deployed in data center ToR or end-of-row applications, as well as for supporting Gigabit Ethernet and 10 Gigabit Ethernet server access in data center end-of-row deployments. During the Lippis/Ixia test, the EX8216 was populated with 128 10GbE ports, classifying it as a Core/spine switch for private or public data center cloud implementations.

Video feature: Click to view Juniper video podcast

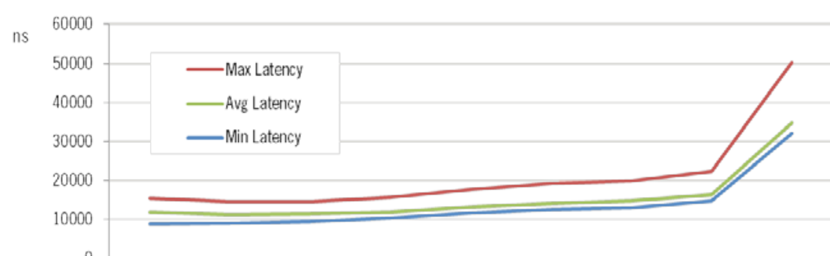
The Juniper EX8216 was tested across all 128 ports of 10GbE. Its average latency ranged from a low of 11366 ns or 11.3 **µs** to a high of 34,723 ns or 34 µs at jumbo size 9216 Byte size frames for layer 2 traffic. Its average delay variation was ranged between 5 and 9.9 ns, providing consistent latency across all packet sizes at full line rate.

For layer 3 traffic, the Juniper EX8216's measured average latency ranged from a low of 11,814 ns to 35,500 ns or 35µs at jumbo size 9216 Byte size frames. Its average delay variation for layer 3 traffic ranged between 5 and 10 ns, providing consistent latency across all packet sizes at full line rate.

Juniper Networks EX8216 Test Configuration

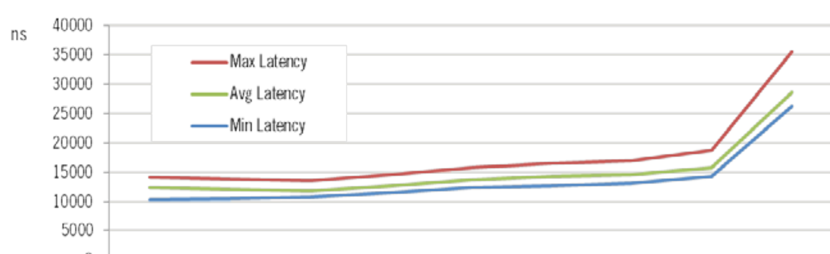
	Hardware	Software Version	Port Density
Device under test	EX8216, EX8200-8XS, EX8200-40XS http://www.juniper.net/us/en/	JUNOS 10.3R2.11	128
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Juniper Networks EX8216 RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	15520	14560	14540	15780	17760	19180	19820	22380	50340
Avg Latency	11891.617	11366.211	11439.586	11966.75	13322.914	14131.773	14789.578	16302.711	34722.688
Min Latency	8860	9100	9500	10380	11800	12480	12980	14720	32100
Avg. delay Variation	9	5	5.375	7.945	7	5.141	9.516	5	9.938

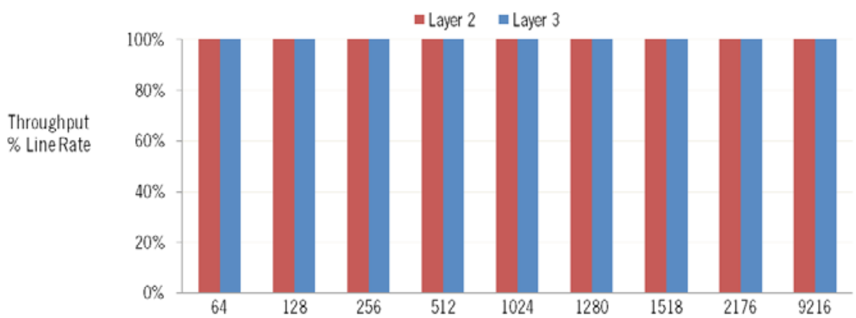
Juniper Networks EX8216 RFC2544 Layer 3 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	14120	13820	13500	14540	15840	16540	17020	18660	35500
Avg Latency	12390.375	12112.5	11814.047	12685.406	13713.25	14318.484	14664.844	15767.438	28674.078
Min Latency	10320	10460	10780	11420	12420	12720	13160	14280	26220
Avg. delay Variation	9	5	5.109	8	7	5	9.563	5	10

 © Lippis Enterprises, Inc. 2011

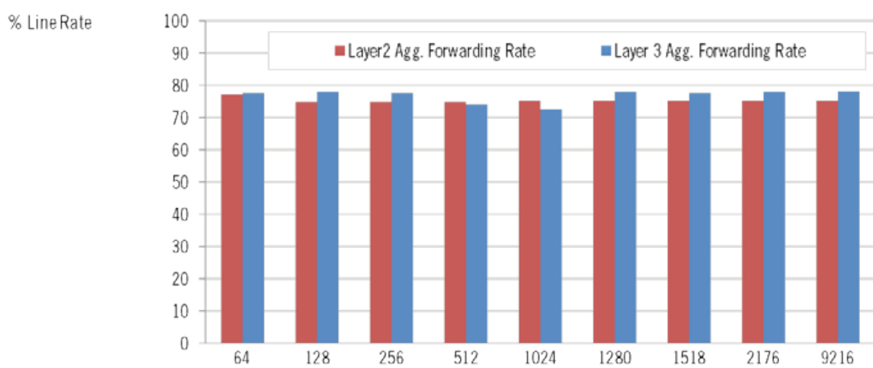
Juniper Networks EX8216
RFC 2544 L2 & L3 Throughput Test



	64	128	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100

The Juniper EX8216 demonstrated 100% throughput as a percentage of line rate across all 128-10GbE ports. In other words, not a single packet was dropped while the Juniper EX8216 was presented with enough traffic to populate its 128 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

Juniper Networks EX8216
RFC 2889 Congestion Test

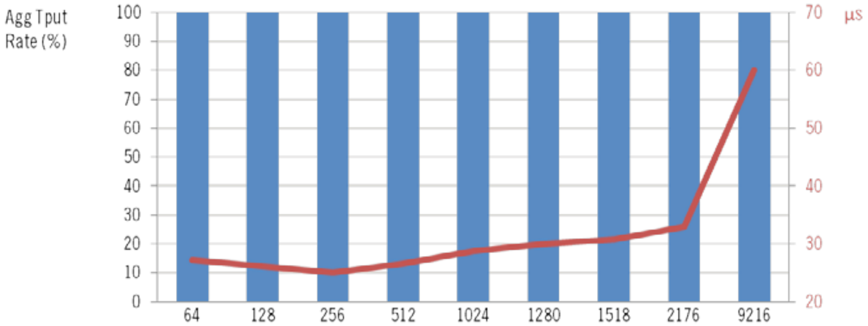


	64	128	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	77.005	75.009	75.016	75.031	75.06	75.06	75.124	75.06	75.06
Layer 3 Agg. Forwarding Rate	77.814	77.835	77.825	74.196	72.784	77.866	77.691	77.91	78.157
Head of Line Blocking	no	no	no	no	no	no	no	no	no
Back Pressure	no	no	no	no	no	no	no	no	no
Agg Flow Control Frames	0	0	0	0	0	0	0	0	0

The Juniper EX8216 demonstrated nearly 80% of aggregated forwarding rate as percentage of line rate during congestion conditions for L2 and L3 forwarding. A single 10GbE port was flooded at 150% of line rate. The EX8216 did not use HOL blocking, which means that as the 10GbE port on the EX8216 became congested, it did not impact the performance of other ports. There was no back pressure detected, and the Ixia test gear did not receive flow control frames.

© Lippis Enterprises, Inc. 2011

Juniper Networks EX8216 RFC 3918 IP Multicast Test



	64	128	256	512	1024	1280	1518	2176	9216
Agg Tput Rate (%)	100	100	100	100	100	100	100	100	100
Agg. Average Latency	27.2	26.2	25.1	26.6	28.8	30.0	30.7	33.0	60.1

The Juniper EX8216 demonstrated 100% and at times 99.99% aggregated throughput for IP multicast traffic with latencies ranging from a 25,135 ns or 25 **µs** to 60,055 ns or 60 **µs** at 9216Byte size packets.

Juniper Networks EX8216 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	30263
East-West	Server_to_Database	11614
East-West	HTTP	18833
East-West	iSCSI-Server_to_Storage	19177
East-West	iSCSI-Storage_to_Server	26927
North-South	Client_to_Server	12145
North-South	Server_to_Client	11322

The Juniper EX8216 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 30µs. Of special note is the consistency that the Juniper EX8216 performed while processing a heavy combination of east-west and north-south flows.

Juniper Networks EX8216 Power Consumption Test

Watts$_{ATIS}$/10GbE port	21.68
3-Year Cost/Watts$_{ATIS}$/10GbE	$79.26
Total power cost/3-Year	$10,145.60
3 yr energy cost as a % of list price	1.30%
TEER Value	44
Cooling	Side-to-Side

The Juniper EX8216 represents a new breed of cloud network spine switches with power efficiency being a core value. Its Watts$_{ATIS}$/port is 21.68 and TEER value is 44. Its power cost per 10GbE is estimated at $26.42 per year. The three-year cost to power the EX8216 is estimated at $10,145.60 and represents less than 2% of its list price. The Juniper EX8216 cooling fans flow side-to-side.

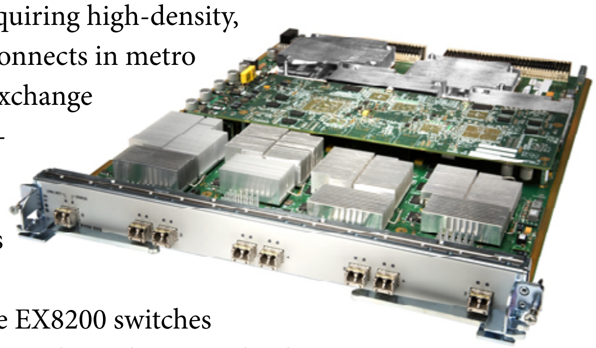
© Lippis Enterprises, Inc. 2011

Discussion:

The EX8216 delivers approximately 1.9 billion packets per second (Bpps) of high-density, wire-speed 10GbE performance for the largest data center networks and includes an advanced set of hardware features enabled by the Juniper-designed EX-PFE2 ASICs. Working with the carrier-class Junos operating system, which runs on all EX Series Ethernet switches, the EX-PFE2 ASICs on each EX8200 line card deliver the scalability to support high-performance data center networks. Two line cards were used in Lippis/Ixia test, the EX8200-8XS and EX8200-40XS described below.

EX8200-8XS Ethernet Line Card

The EX8200-8XS is an 8-port 10GBASE-X line card with compact, modular SFP+ fiber optic interfaces, enabling up to 128 line-rate 10-Gigabit Ethernet ports in an EX8216 modular switch chassis. The EX8200-8XS is ideal for enterprise applications such as campus or data center uplink aggregation, core and backbone interconnects, and for service provider deployments requiring high-density, wire-speed 10GbE interconnects in metro area networks, Internet exchange points and points of presence (POPs).

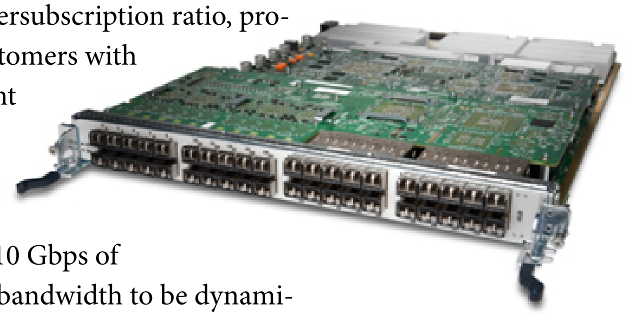


The 10GbE port densities afforded by the EX8200-8XS line cards also enable EX8200 switches to consolidate aggregation and core layers in the data center, simplifying the network architecture and reducing power, space and cooling requirements while lowering total cost of ownership (TCO).

EX8200-40XS Ethernet Line Card

The EX8200-40XS is a 40-port oversubscribed 10GbE solution for data center end-of-row and middle-of-row server access, as well as for data center blade switch and top-of-rack or campus uplink aggregation deployments. The EX8200-40XS supports a wide range of both SFP (GbE) and SFP+ (10 GbE) modular optical interfaces for connecting over multimode fiber, single-mode fiber, and copper cabling.

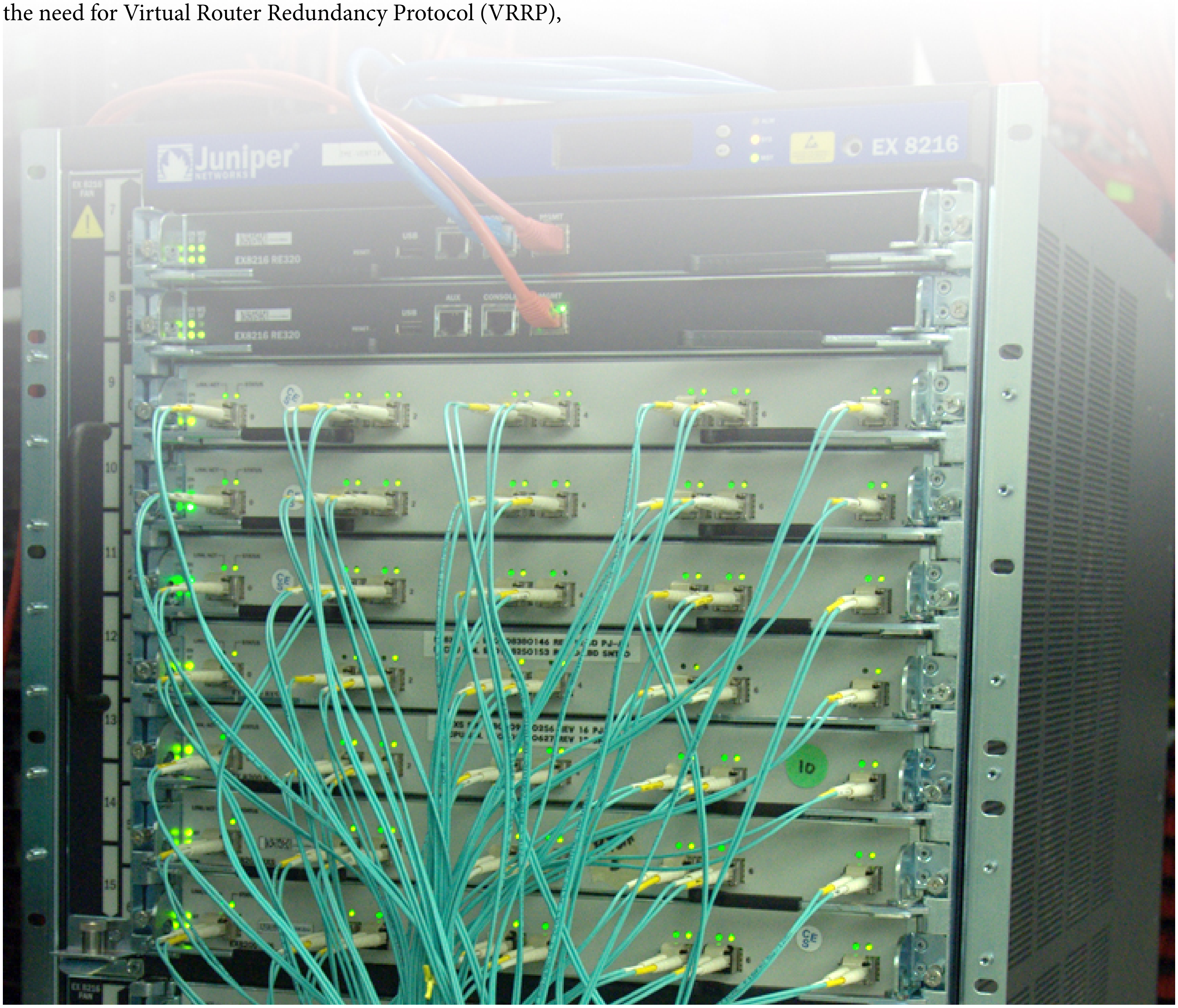
The 40 SFP/SFP+ ports on the EX8200-40XS are divided into 8 independent groups of 5 ports each. Because port groups are independent of one another, each group can have its own oversubscription ratio, providing customers with deployment flexibility. Each group dedicates 10 Gbps of switching bandwidth to be dynamically shared among the ports; queues are allocated within a 1MB oversubscription buffer based on the number of active ports in a group and the types of interfaces installed. Users simply connect the cables and the EX8200-40XS automatically provisions each port group accordingly. No manual configuration is required.



© Lippis Enterprises, Inc. 2011

Virtual Chassis Technology

The EX8200 supports Juniper Networks' unique Virtual Chassis technology, which enables two interconnected chassis—any combination of EX8208s or EX8216s—to operate as a single, logical device with a single IP address. Deployed as a collapsed aggregation or core layer solution, an EX8200 Virtual Chassis configuration creates a network fabric for interconnecting access switches, routers and service-layer devices, such as firewalls and load balancers, using standards-based Ethernet LAGs. The network fabric created by an EX8200 Virtual Chassis configuration prevents loops, eliminating the need for protocols such as Spanning Tree. The fabric also simplifies the network by eliminating the need for Virtual Router Redundancy Protocol (VRRP), increasing the scalability of the network design. EX8200 Virtual Chassis configurations are highly resilient, with no single point of failure, ensuring that no single element can render the entire fabric inoperable following a failure. In an EX8200 Virtual Chassis configuration, the Routing Engine functionality is externalized to a purpose-built, server-class appliance, the XRE200, which supports control plane processing requirements for large-scale systems and provides an extra layer of availability and redundancy. All control protocols such as OSPF, IGMP, Link Aggregation Control Protocol (LACP), 802.3ah and VCCP, as well as all management plane functions, run or reside on the XRE200.



© Lippis Enterprises, Inc. 2011

Voltaire® Vantage™ 6048

The Voltaire Vantage™ 6048 switch is a high performance Layer 2 10GbE ToR/leaf switch optimized for enterprise data center and cloud computing environments, with 48 ports of 10GbE line-rate connectivity and 960 Gb/s non-blocking switching throughput.

Video feature: Click to view Voltaire video podcast

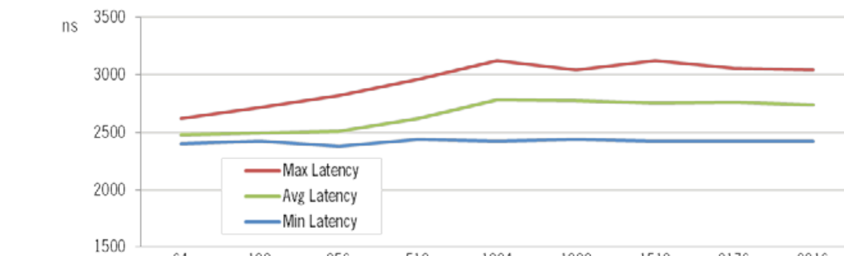
The Voltaire Vantage™ 6048 ToR switch was tested across all 48 ports of 10GbE. Its average latency ranged from a low of 2484 ns to a high of 2784 ns for layer 2 traffic. Its average delay variation ranged between 4.8 and 9.6 ns, providing consistent latency across all packet sizes at full line rate.

For layer 3 traffic, the Voltaire Vantage™ 6048's average latency ranged from a low of 2504 to 2791 ns across all frame sizes. Its average delay variation for layer 3 traffic ranged between 5 and 9.5 ns, providing consistent latency across all packet sizes at full line rate.

Voltaire Vantage 6048 Test Configuration

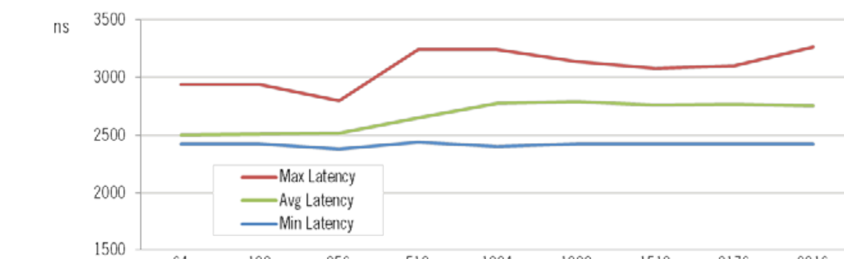
	Hardware	Software Version	Port Density
Device under test	6048 http://www.voltaire.com	1.0.0.50	48
Test Equipment	Ixia XM12 High Performance Chassis	IxOS 6.00 EA	
	Xcellon Flex AP10G16S 10 Gigabit Ethernet LAN load module	IxNetwork 5.70 EA	
	LSM10GXM8S and LSM10GXM8XP 10 Gigabit Ethernet load modules		
	1U Application Server	IxAutomate 6.90 GA SP1	
	http://www.ixiacom.com/		
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SPF+ transceivers		
	www.leviton.com		

Voltaire® Vantage™ 6048 RFC2544 Layer 2 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	2620	2720	2820	2960	3120	3040	3120	3060	3040
Avg Latency	2484.354	2495.521	2513.521	2617.563	2783.833	2776.146	2757.021	2759.813	2737.354
Min Latency	2400	2420	2380	2440	2420	2440	2420	2420	2420
Avg. delay Variation	8.625	5.208	5.146	7.479	6.979	4.875	9.688	4.958	9.417

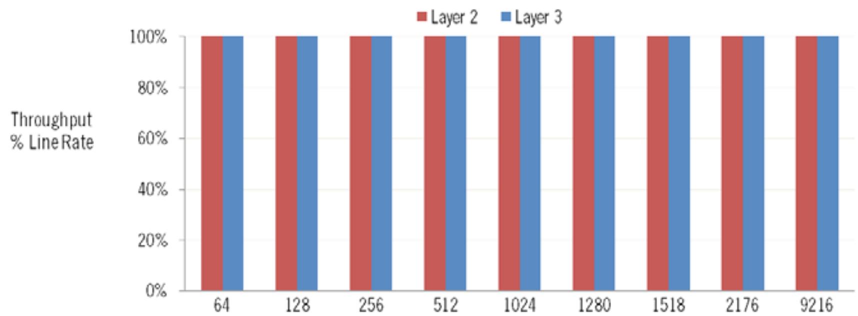
Voltaire® Vantage™ 6048 RFC2544 Layer 3 Latency Test



	64	128	256	512	1024	1280	1518	2176	9216
Max Latency	2940	2940	2800	3240	3240	3140	3080	3100	3260
Avg Latency	2504.188	2513.625	2517.979	2650.021	2773.729	2791.313	2762.813	2766.958	2753.958
Min Latency	2420	2420	2380	2440	2400	2420	2420	2420	2420
Avg. delay Variation	8.583	5.313	5.188	7.667	6.958	5.479	9.583	5	9.313

 © Lippis Enterprises, Inc. 2011

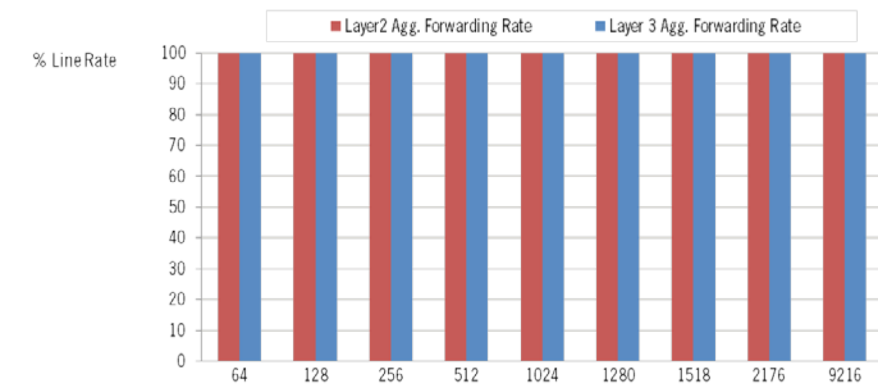
Voltaire® Vantage™ 6048
RFC 2544 L2 & L3 Throughput Test



	64	128	256	512	1024	1280	1518	2176	9216
Layer 2	100	100	100	100	100	100	100	100	100
Layer 3	100	100	100	100	100	100	100	100	100

The Voltaire Vantage™ 6048 demonstrated 100% throughput as a percentage of line rate across all 48-10GbE ports. In other words, not a single packet was dropped while the Voltaire Vantage™ 6048 was presented with enough traffic to populate its 48 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

Voltaire® Vantage™ 6048
RFC 2889 Congestion Test



	64	128	256	512	1024	1280	1518	2176	9216
Layer2 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Layer 3 Agg. Forwarding Rate	100	100	100	100	100	100	100	100	100
Head of Line Blocking	no	no	no	no	no	no	no	no	no
Back Pressure	yes	yes	yes	yes	yes	yes	yes	yes	yes
L2 Agg Flow Control Frames	7152760	5201318	2711124	1721416	1756370	2110202	1807994	2168202	2079506
L3 Agg Flow Control Frames	7152846	5215032	2710746	1721416	1756238	2110308	1808060	2172256	2079658

The Voltaire Vantage™ 6048 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions for all packet sizes during the L2 and L3 forwarding test. A single 10GbE port was flooded at 150% of line rate. The Vantage™ 6048 did not use HOL blocking, which means that as the 10GbE port on the Vantage™ 6048 became congested, it did not impact the performance of other ports. Back pressure was present throughout all packets sizes as the Voltaire Vantage™ 6048 sent flow control frames to the Ixia test gear, which is normal operation.

For layer 3-traffic, the Vantage™ 6048 demonstrated 100% of aggregated forwarding rate as percentage of line rate during congestion conditions for all packet sizes. No HOL blocking was present. Back pressure was present throughout all packets sizes as the Voltaire Vantage™ 6048 sent flow control frames to the Ixia test gear, which is normal operation.

© Lippis Enterprises, Inc. 2011

Voltaire® Vantage™ 6048
RFC 3918 IP Multicast Test

The Voltaire Vantage™ 6048 does not support IP multicast.

Voltaire Vantage™ 6048 Cloud Simulation Test

Traffic Direction	Traffic Type	Avg Latency (ns)
East-West	Database_to_Server	11401
East-West	Server_to_Database	2348
East-West	HTTP	10134
East-West	iSCSI-Server_to_Storage	2884
East-West	iSCSI-Storage_to_Server	10454
North-South	Client_to_Server	3366
North-South	Server_to_Client	2197

The Voltaire Vantage™ 6048 performed well under cloud simulation conditions by delivering 100% aggregated throughput while processing a large combination of east-west and north-south traffic flows. Zero packet loss was observed as its latency stayed under 11401 ns or 11 µs.

Voltaire Vantage™ 6048 Power Consumption Test

$Watts_{ATIS}$/10GbE port	5.5
3-Year Cost/$Watts_{ATIS}$/10GbE	$20.11
Total power cost/3-Year	$965.19
3 yr energy cost as a % of list price	4.17%
TEER Value	172
Cooling	Front to Back

The Voltaire Vantage™ 6048 represents a new breed of cloud network leaf or ToR switches with power efficiency being a core value. Its $Watts_{ATIS}$/port is 5.5 and TEER value is 172. Note higher TEER values are more desirable as they represent the ratio of work performed over energy consumption. Its power cost per 10GbE is estimated at $6.70 per year. The three-year cost to power the Voltaire Vantage™ 6048 is estimated at $965.19 and represents approximately 4% of its list price. Keeping with data center best practices, its cooling fans flow air front to back.

© Lippis Enterprises, Inc. 2011

Discussion

The Voltaire Vantage™ 6048 is a 48-port 10GbE ToR switch. Not tested in the Lippis/Ixia test are its claimed converged Ethernet capabilities to enable new levels of efficiency, scalability and real-time application performance, while at the same time consolidating multiple/redundant network tiers and significantly reducing infrastructure expenses.

 © Lippis Enterprises, Inc. 2011

Top-of-Rack Switches Cross-Vendor Analysis

Between the two-industry Lippis/Ixia tests of December 2010 and April of 2011 there were eight ToR switches evaluated for performance and power consumption. These participating vendors are:

Arista 7124SX 10G SFP Data Center Switch

Arista 7050S-64 10/40G Data Center Switch

BLADE Network Technologies, an IBM Company IBM BNT RackSwitch G8124

BLADE Network Technologies, an IBM Company IBM BNT RackSwitch G8264

Brocade VDX™ 6720-24 Data Center Switch

Force10 S-Series S4810

Hitachi Cable Apresia 15000-64XL-PSR

Voltaire® Vantage™ 6048

All but Brocade utilize merchant silicon in their ToR switch. The rest utilize a new single chip design from Broadcom, Marvel or Fulcrum Microsystems. Brocade developed its own single chip silicon. With a single chip provided by a chip manufacturer, vendors are free to invest resources other than ASIC development, which can consume much of a company's engineering and financial resources. With merchant silicon providing a forwarding engine for their switches, these vendors are free to choose where to innovate, be it in buffer architecture, network services such as virtualization support, 40GbE uplink or fan-in support, etc. The Lippis/Ixia test results demonstrate that these new chip designs provide state-of-the-art performance at efficient power consumption levels not seen before. Further, price points on a 10GbE per port basis are a low of $351 to a high of $670.

IT business leaders are responding favorably to ToR switches equipped with a value proposition of high performance, low acquisition price and low power consumption. These ToR switches currently are the hot boxes in the industry with quarterly revenues for mid-size firms in the $10 to $15M plus. We compared each of the above firms in terms of their ability to forward packets: quickly (i.e., latency), without loss of their throughput at full line rate, when ports are oversubscribed with network traffic by 150 percent, in IP multicast mode and in cloud simulation. We also measure their power consumption as described in the Lippis Test Methodology section above.

In the following ToR analysis we separate 24 port ToR switches from all others. There are three 24 port ToR switches:

Arista 7124SX 10G SFP Data Center Switch

BLADE Network Technologies, an IBM Company IBM BNT RackSwitch G8124

Brocade VDX™ 6720-24 Data Center Switch

The 48 port and above ToR switches evaluated are:

Arista 7050S-64 10/40G Data Center Switch

BLADE Network Technologies, an IBM Company IBM BNT RackSwitch G8264

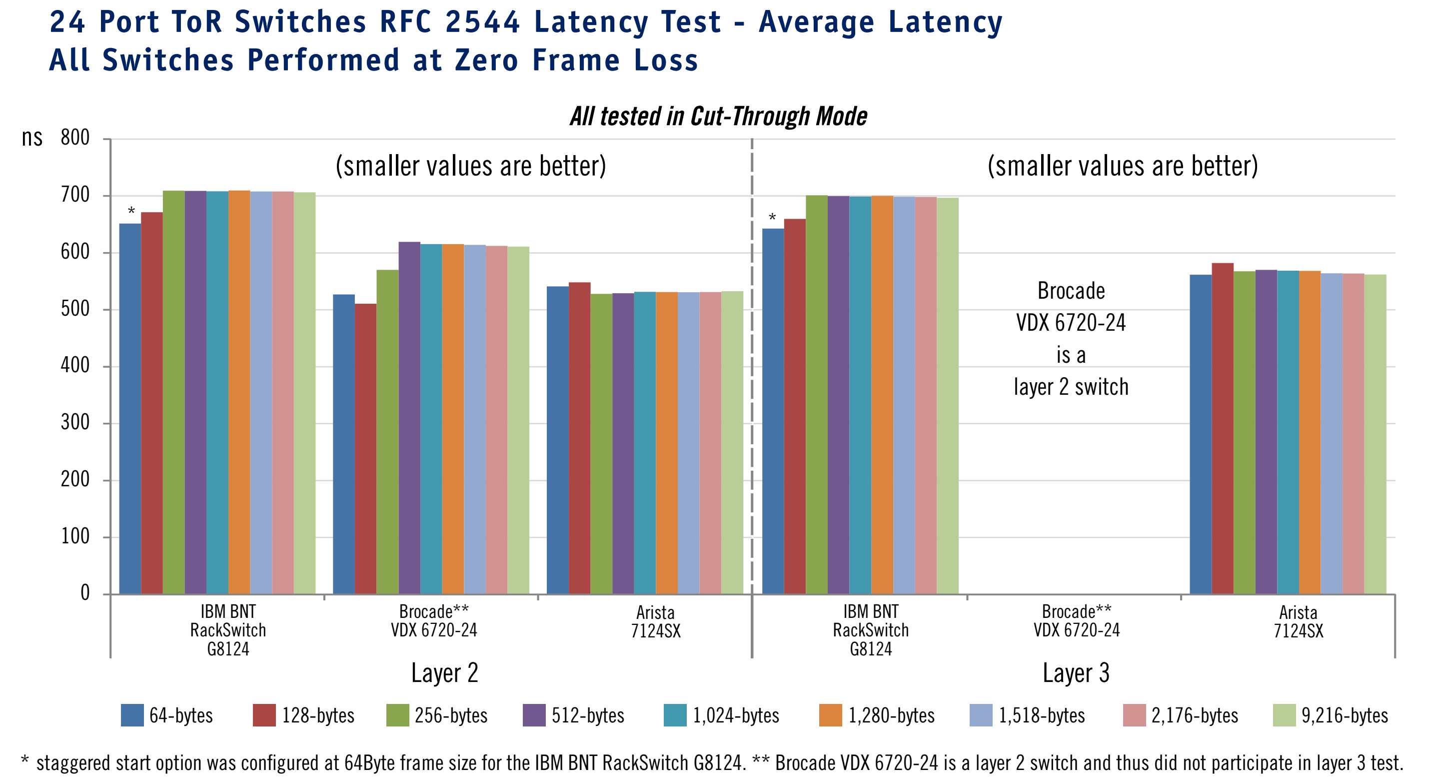
Force10 S-Series S4810

Hitachi Cable Apresia 15000-64XL-PSR

Voltaire. Vantage™ 6048

This approach provides a consistent cross-vendor analysis. We start with a cross-vendor analysis of 24 Port ToR switches followed by the 48 port and above devices and with a set of industry recommendations.

© Lippis Enterprises, Inc. 2011



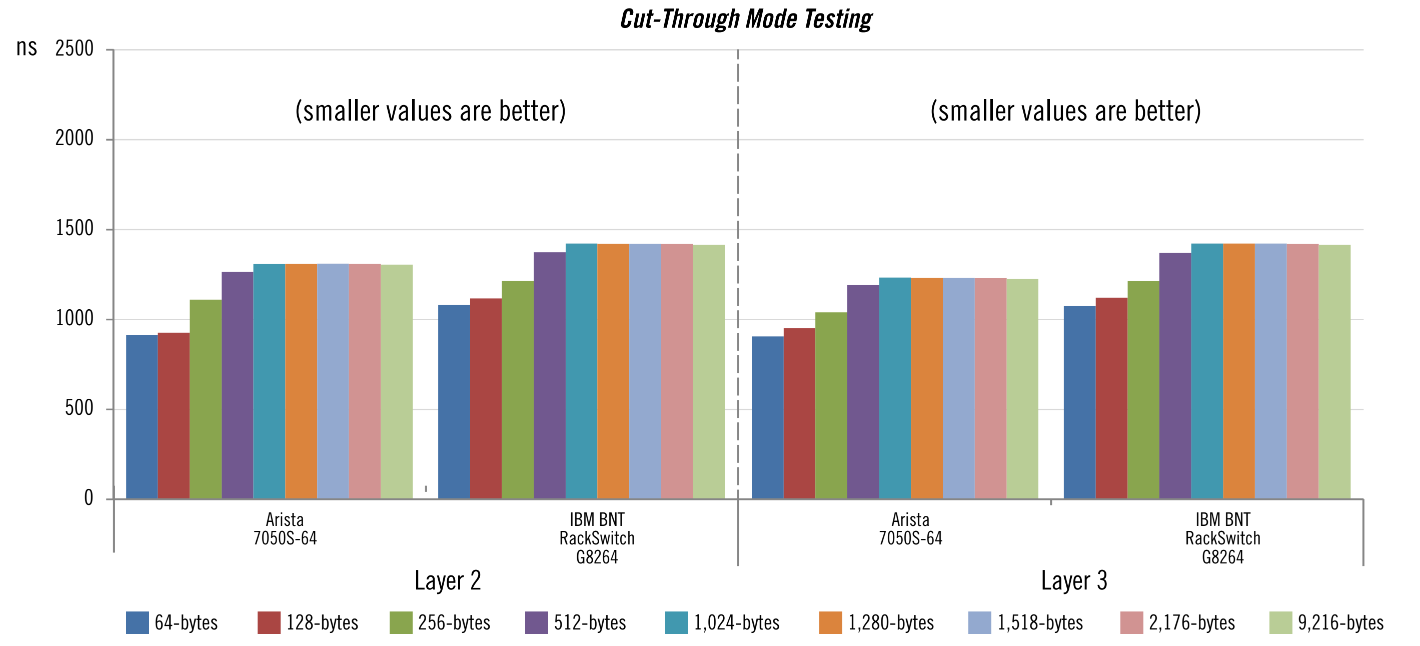
	Layer 2			Layer 3		
Frames/Packet Size (bytes)	IBM BNT* RackSwitch G8124	Brocade VDX 6720-24	Arista 7124SX	IBM BNT* RackSwitch G8124	Brocade VDX 6720-24	Arista 7124SX
64	651.75	527.04	541.08	642.71	**	561.58
128	671.50	510.71	548.21	659.67	**	582.33
256	709.46	570.08	527.92	701.29	**	567.75
512	709.04	619.46	528.96	699.79	**	570.29
1,024	708.25	615.54	531.58	699.00	**	568.79
1,280	709.71	615.58	531.42	699.92	**	568.54
1,518	707.93	614.21	531.04	698.50	**	564.33
2,176	708.04	612.17	531.42	698.33	**	563.83
9,216	706.63	610.96	532.79	696.92	**	562.21

* staggered start option was configured at 64Byte frame size for the IBM BNT RackSwitch G8124. ** Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test.

We show average latency and average delay variation across all packet sizes for layer 2 and 3 forwarding. Measurements were taken from ports that were far away from each other to demonstrate the consistency of performance across the single chip design. The stand out here is that these switch latencies are as low as 510ns to a high of 709ns. The Arista 7124SX layer 2 latency is the lowest measured in the Lippis/Ixia set of cloud networking test. Also of note is the consistency in low latency across all packet sizes.

© Lippis Enterprises, Inc. 2011

ToR Switches RFC 2544 Latency Test - Average Latency
All Switches Performed at Zero Frame Loss



Cut-Through Mode Testing

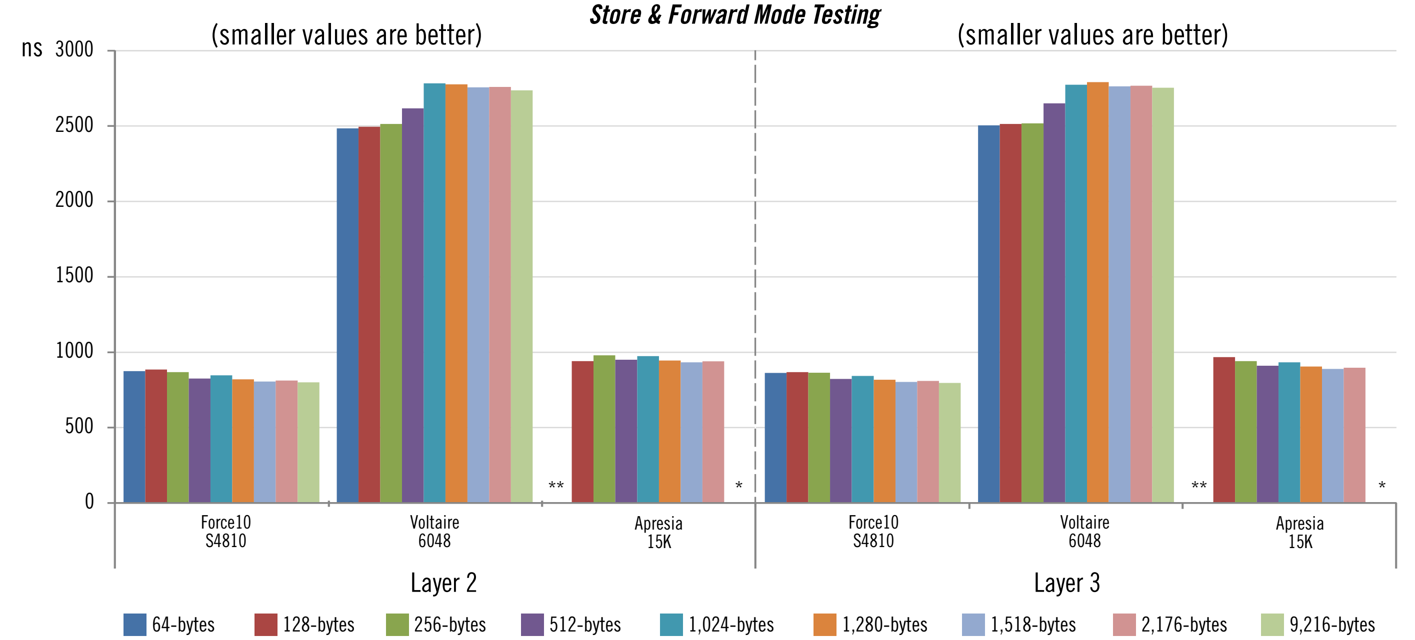
Frames/Packet Size (bytes)	Layer 2: Arista 7050S-64	Layer 2: IBM BNT RackSwitch G8264	Layer 3: Arista 7050S-64	Layer 3: IBM BNT RackSwitch G8264
64	914.33	1081.34	905.42	1075.39
128	926.77	1116.83	951.20	1121.02
256	1110.64	1214.92	1039.14	1213.55
512	1264.91	1374.25	1191.64	1370.94
1,024	1308.84	1422.00	1233.00	1421.94
1,280	1309.11	1421.28	1232.45	1422.63
1,518	1310.77	1421.75	1232.48	1421.83
2,176	1309.77	1419.89	1229.92	1419.73
9,216	1305.50	1415.36	1225.58	1415.78

The Arista 7124SX, 7050S-64, IBM BNT G8264, G8124 and Brocade VDX™ 6720-24 switches were configured and tested via cut-through test method, while Force10, Voltaire and Apresia were configured and tested via store-and-forward method. This is due to the fact that some switches are store and forward while others are cut through devices. During store-and-forward testing, the latency of packet serialization delay (based on packet size) is removed from the reported latency number, by test equipment. Therefore, to compare cut-through to store-and-forward latency measurement, packet serialization delay needs to be added to store-and-forward latency number. For example, in a store-and-forward latency number of 800ns for a 1,518 byte size packet, the additional latency of 1240ns (serialization delay of a 1518 byte packet at 10Gbps) is required to be added to the store-and-forward measurement. This difference can be significant. Note that other potential device specific factors can impact latency too. This makes comparisons between two testing methodologies difficult.

As in the 24 port ToR switches, we show average latency and average delay variation across all packet sizes for layer 2 and 3 forwarding. Measurements were taken from ports that were far away from each other to demonstrate the consistency of performance across the single chip design. These new generation of switches support both cut-through and store and forward forwarding, so we separate these results. Both the Arista 7050S-64 offers approximately 100ns less latency than the IBM BNT RackSwitch G8264.

© Lippis Enterprises, Inc. 2011

ToR Switches RFC 2544 Latency Test - Average Latency
All Switches Performed at Zero Frame Loss



* largest frame size supported on Apresia 15K is 9044. ** There is no latency data for the Apresia 15K at 64 bytes due to configuration difficulties during testng. The 15K could not be configured to maintain a VLAN @ 64 Bytes which eliminated packet signature to measure latency.

The Voltaire 6048 measured latency is the highest for store and forward devices while Force10 S4810 was the lowest across nearly all packet sizes.

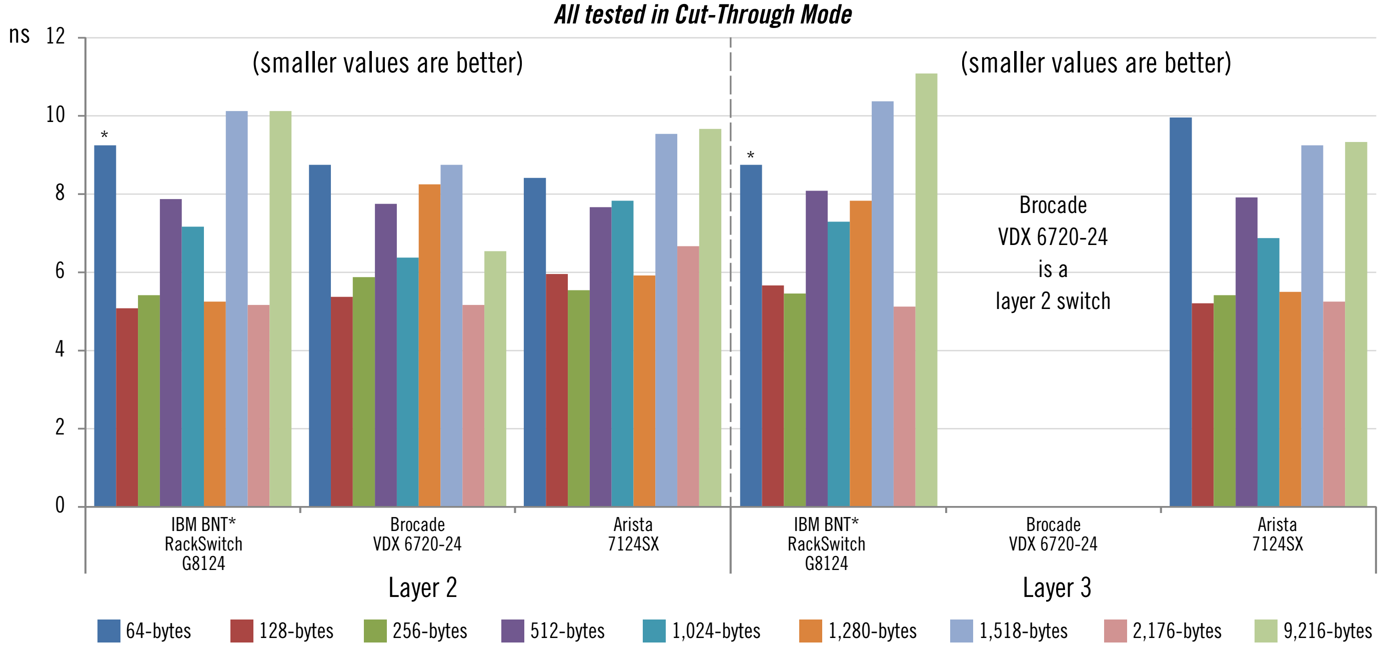
Store & Forward Mode Testing

	Layer 2			Layer 3		
Frames/Packet Size (bytes)	Force10 S4810	Voltaire 6048	Apresia 15K	Force10 S4810	Voltaire 6048	Apresia 15K
64	874.48	2484.35	**	862.42	2504.19	**
128	885.00	2495.52	941	868.54	2513.63	967.17
256	867.60	2513.52	980	863.77	2517.98	941.08
512	825.29	2617.56	950	822.63	2650.02	910.84
1,024	846.52	2783.83	974	843.02	2773.73	933.05
1,280	819.90	2776.15	945	817.83	2791.31	904.84
1,518	805.06	2757.02	933	802.38	2762.81	889.55
2,176	812.06	2759.81	940	809.71	2766.96	897.33
9,216	799.79	2737.35	*900.64	796.06	2753.96	*900.64

* largest frame size supported on Apresia 15K is 9044. ** There is no latency data for the Apresia 15K at 64 bytes due to configuration difficulties during testng. The 15K could not be configured to maintain a VLAN @ 64 Bytes which eliminated packet signature to measure latency.

© Lippis Enterprises, Inc. 2011

24 Port ToR Switches RFC 2544 Latency Test - Average Delay Variation
All Switches Performed at Zero Frame Loss



* staggered start option was configured at 64Byte frame size for the IBM BNT RackSwitch G8124. ** Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test

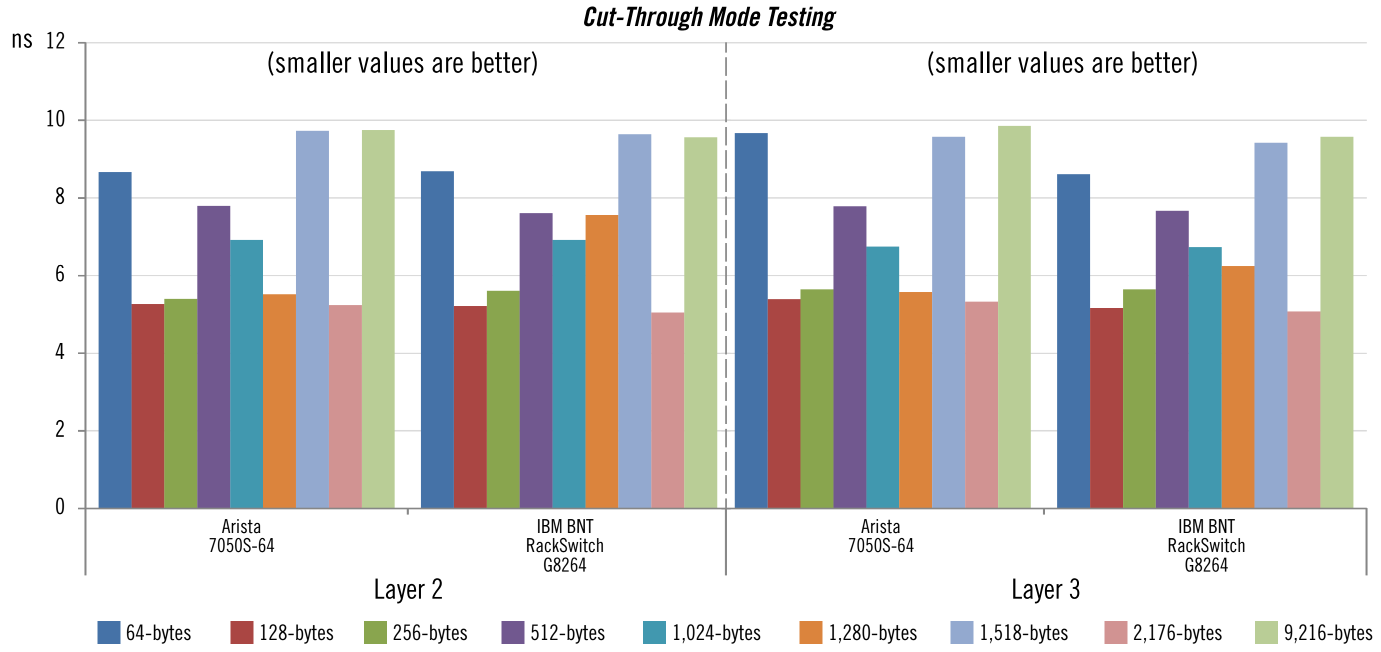
	Layer 2			Layer 3		
Frames/Packet Size (bytes)	IBM BNT RackSwitch G8124**	Brocade VDX 6720-24	Arista 7124SX	IBM BNT RackSwitch G8124**	Brocade VDX 6720-24	Arista 7124SX
64	9.25	8.75	8.42	8.75	**	9.96
128	5.08	5.38	5.96	5.67	**	5.21
256	5.42	5.88	5.54	5.46	**	5.42
512	7.88	7.75	7.67	8.09	**	7.92
1,024	7.17	6.38	7.83	7.30	**	6.88
1,280	5.25	8.25	5.92	7.84	**	5.50
1,518	10.13	8.75	9.54	10.38	**	9.25
2,176	5.17	5.17	6.67	5.13	**	5.25
9,216	10.13	6.54	9.67	11.08	**	9.33

* staggered start option was configured at 64Byte frame size for the IBM BNT RackSwitch G8124. ** Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test

Notice there is little difference in average delay variation across all vendors proving consistent latency under heavy load at zero packet loss for L2 and L3 forwarding. Difference does reside within average latency between suppliers. The range of average delay variation is 5ns to 10ns.

© Lippis Enterprises, Inc. 2011

ToR Switches RFC 2544 Latency Test - Average Delay Variation
All Switches Performed at Zero Frame Loss



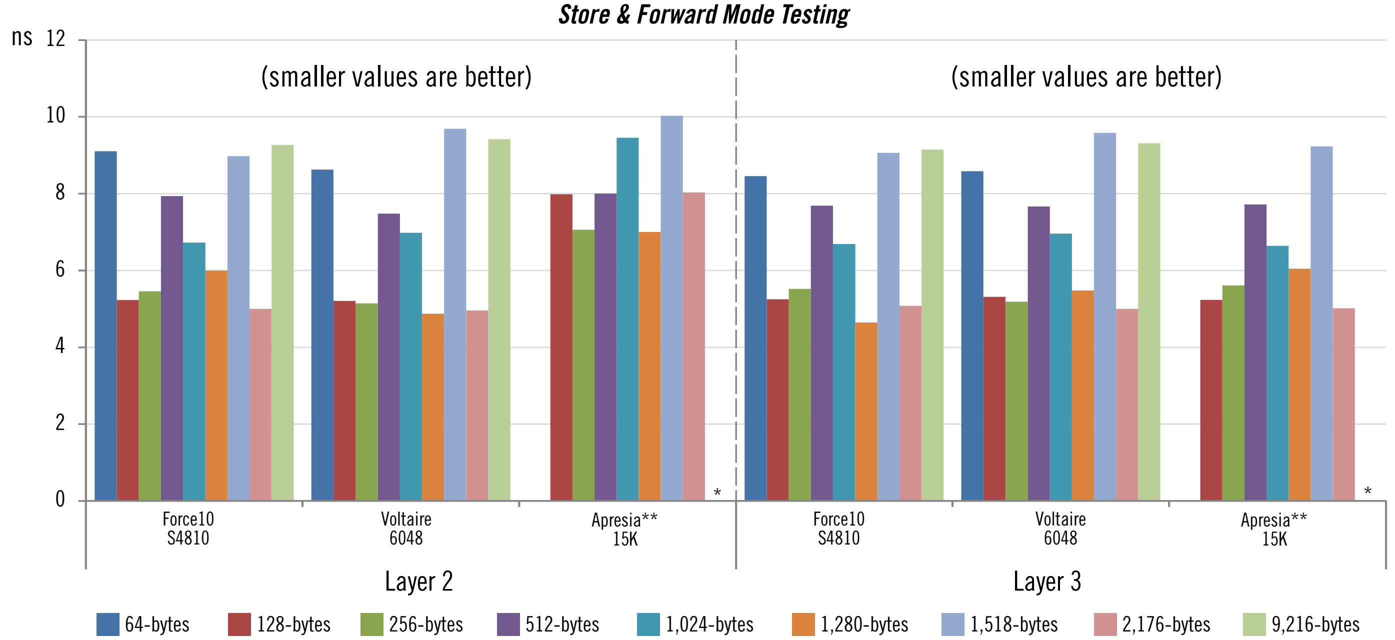
As in the 24 port ToR switches, the 48 port and above ToR devices show slight difference in average delay variation between Arista 7050S-64 and IBM BNT RackSwitch G8264 thus proving consistent latency under heavy load at zero packet loss for Layer 2 and Layer 3 forwarding. Difference does reside within average latency between suppliers. Just as in the 24 port ToR switches, the range of average delay variation is 5ns to 10ns.

Cut-Through Mode Testing

	Layer 2		Layer 3	
Frames/Packet Size (bytes)	Arista 7050S-64	IBM BNT RackSwitch G8264	Arista 7050S-64	IBM BNT RackSwitch G8264
64	8.67	8.69	9.67	8.61
128	5.27	5.22	5.39	5.17
256	5.41	5.61	5.64	5.64
512	7.80	7.61	7.78	7.67
1,024	6.92	6.92	6.75	6.73
1,280	5.52	7.56	5.58	6.25
1,518	9.73	9.64	9.58	9.42
2,176	5.23	5.05	5.33	5.08
9,216	9.75	9.56	9.86	9.58

© Lippis Enterprises, Inc. 2011

ToR Switches RFC 2544 Latency Test - Average Delay Variation
All Switches Performed at Zero Frame Loss



* largest frame size supported on Apresia 15K is 9044. There is no latency data for the Apresia 15K at 64 bytes due to configuration difficulties during testng. The 15K could not be configured to maintain a VLAN @ 64 Bytes which eliminated packet signature to measure latency.

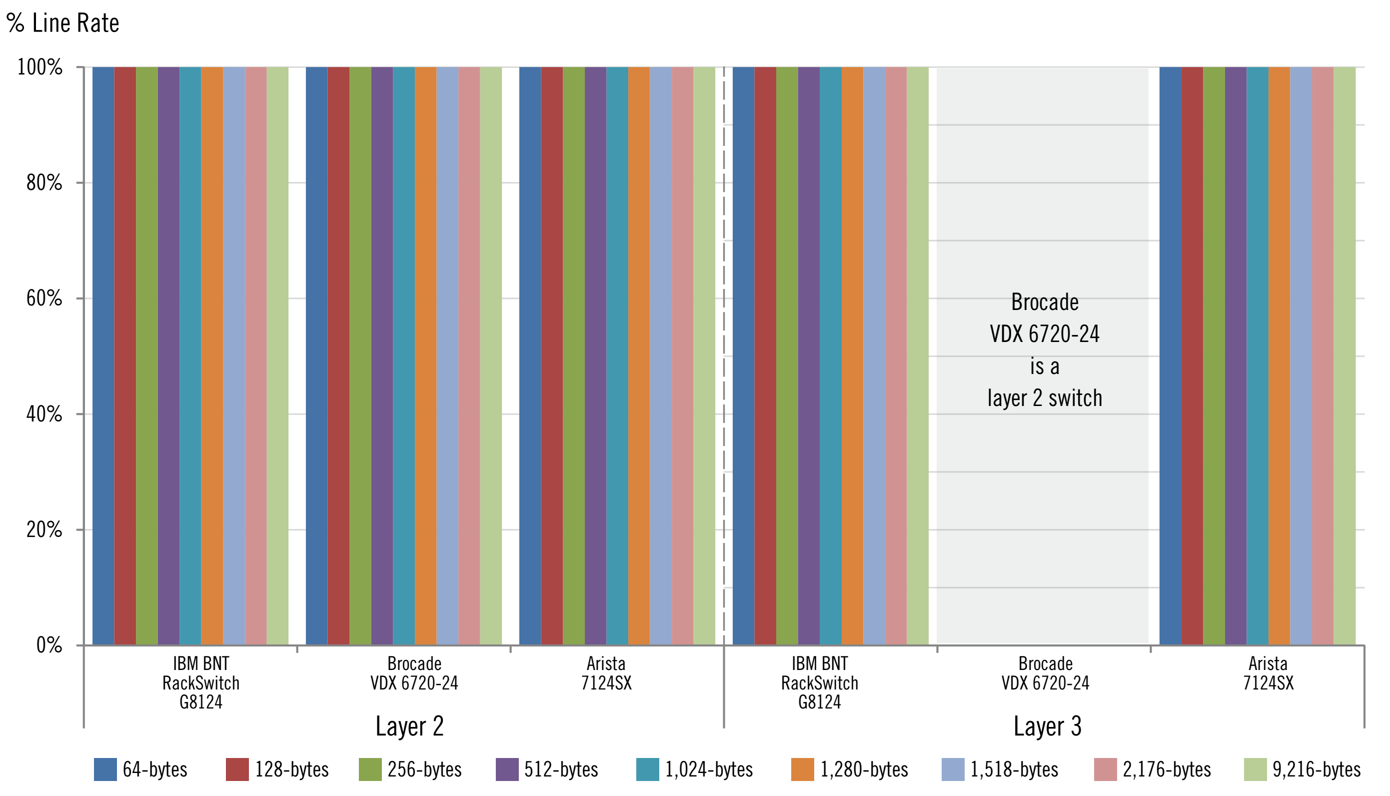
Store & Forward Mode Testing

	Layer 2			Layer 3		
Frames/Packet Size (bytes)	Force10 S4810	Voltaire 6048	Apresia 15K	Force10 S4810	Voltaire 6048	Apresia 15K
64	9.10	8.63	**	8.46	8.58	**
128	5.23	5.21	8.00	5.25	5.31	5.23
256	5.46	5.15	7.00	5.52	5.19	5.61
512	7.94	7.48	8.00	7.69	7.67	7.72
1,024	6.73	6.98	9.00	6.69	6.96	6.64
1,280	6.00	4.88	7.00	4.65	5.48	6.05
1,518	8.98	9.69	10.00	9.06	9.58	9.23
2,176	5.00	4.96	8.00	5.08	5.00	5.02
9,216	9.27	9.42	*10.58	9.15	9.31	*9.67

* largest frame size supported on Apresia 15K is 9044. There is no latency data for the Apresia 15K at 64 bytes due to configuration difficulties during testng. The 15K could not be configured to maintain a VLAN @ 64 Bytes which eliminated packet signature to measure latency.

The 48 port and above ToR devices latency were measured in cut-through mode and show slight difference in average delay variation between Force10 S4810, Voltaire VantageTM 6048 and Apresia 15K. All switches demonstrate consistent latency under heavy load at zero packet loss for Layer 2 and Layer 3 forwarding. The Apresia 15K does demonstrate largest average delay variation across all packet sizes while Force10 and Voltaire are more varied. However, the range of average delay variation is 5ns to 10ns.

© Lippis Enterprises, Inc. 2011



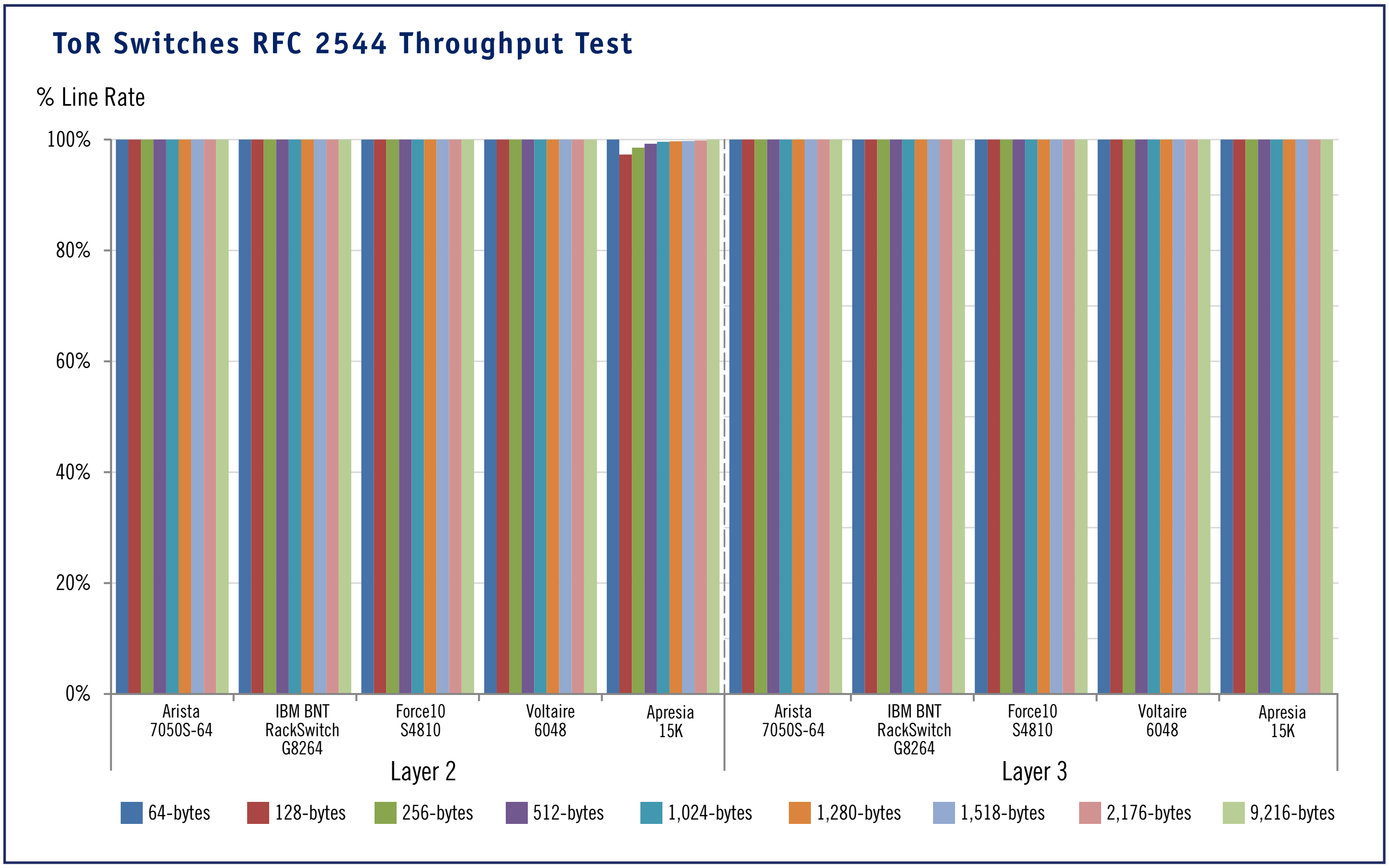
* Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test.

	Layer 2			Layer 3		
Frames/Packet Size (bytes)	IBM BNT RackSwitch G8124	Brocade VDX 6720-24	Arista 7124SX	IBM BNT RackSwitch G8124	Brocade VDX 6720-24	Arista 7124SX
64	100%	100%	100%	100%	*	100%
128	100%	100%	100%	100%	*	100%
256	100%	100%	100%	100%	*	100%
512	100%	100%	100%	100%	*	100%
1,024	100%	100%	100%	100%	*	100%
1,280	100%	100%	100%	100%	*	100%
1,518	100%	100%	100%	100%	*	100%
2,176	100%	100%	100%	100%	*	100%
9,216	100%	100%	100%	100%	*	100%

* Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test.

As expected all 24 port ToR switches are able to forward L2 and L3 packets at all sizes at 100% of line rate with zero packet loss, proving that these switches are high performance wire-speed devices.

© Lippis Enterprises, Inc. 2011

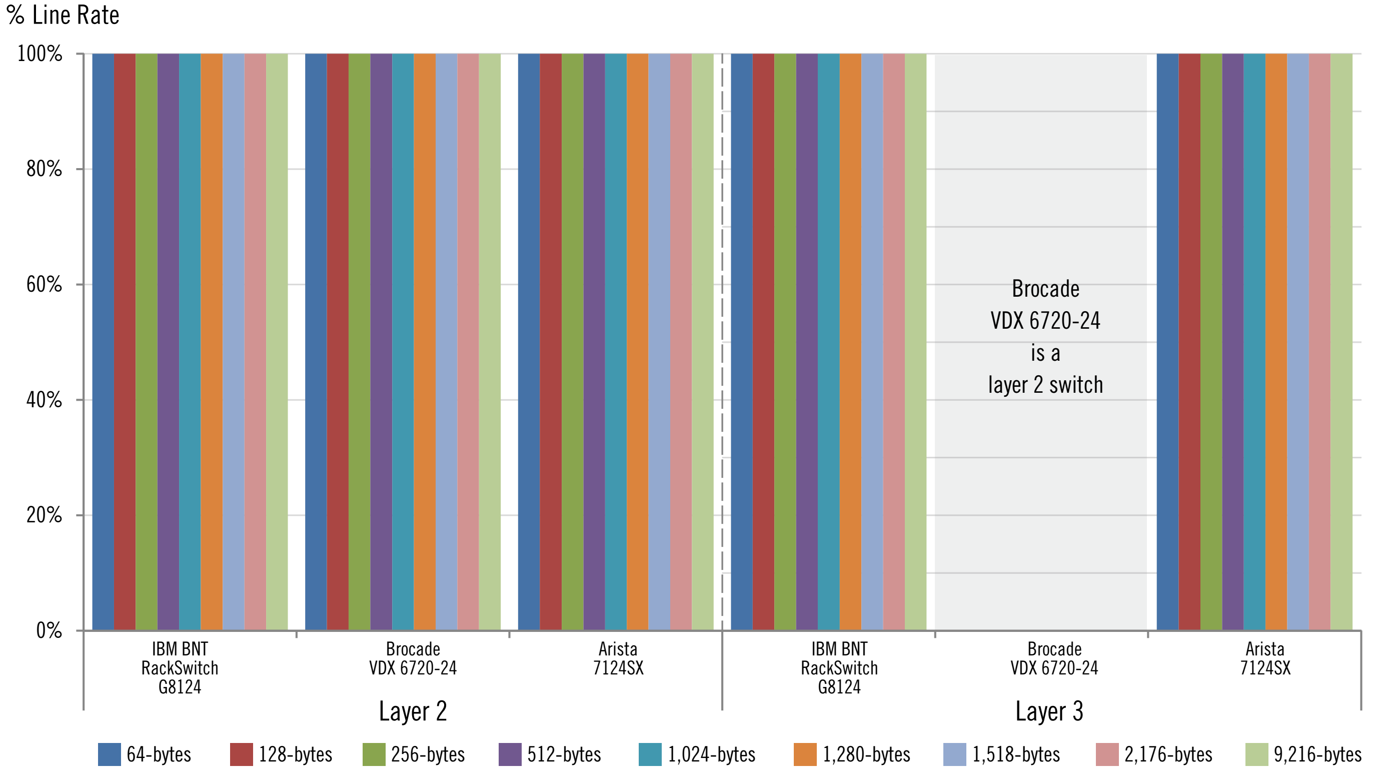


	Layer 2					Layer 3				
Frames/Packet Size (bytes)	Arista 7050S-64	IBM BNT RackSwitch G8264	Force10 S4810	Voltaire 6048	Apresia 15K	Arista 7050S-64	IBM BNT RackSwitch G8264	Force10 S4810	Voltaire 6048	Apresia 15K
64	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
128	100%	100%	100%	100%	97.288%	100%	100%	100%	100%	100%
256	100%	100%	100%	100%	98.541%	100%	100%	100%	100%	100%
512	100%	100%	100%	100%	99.238%	100%	100%	100%	100%	100%
1,024	100%	100%	100%	100%	99.607%	100%	100%	100%	100%	100%
1,280	100%	100%	100%	100%	99.682%	100%	100%	100%	100%	100%
1,518	100%	100%	100%	100%	99.73%	100%	100%	100%	100%	100%
2,176	100%	100%	100%	100%	99.808%	100%	100%	100%	100%	100%
9,216	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

With the exception of the Apresia 15K, as expected all switches are able to forward L2 and L3 packets at all sizes at 100% of line rate with zero packet loss, proving that these switches are high performance wire-speed devices.

© Lippis Enterprises, Inc. 2011

24 Port ToR Switches RFC 2889 Congestion Test
150% of line rate into single 10GbE



	Layer 2									Layer 3								
Frames/ Packet Size (bytes)	IBM BNT RackSwitch G8124			Brocade VDX 6720-24			Arista 7124SX			IBM BNT RackSwitch G8124			Brocade VDX 6720-24			Arista 7124SX		
	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP
64	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
128	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
256	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
512	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
1,024	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
1,280	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
1,518	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
2,176	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y
9,216	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*	N	Y	100%	N	Y

* Brocade VDX 6720-24 is a layer 2 switch and thus did not participate in layer 3 test. AFR% = Agg Forwarding Rate (% Line Rate) HOL = Head of Line Blocking BP = Back Pressure
Information was recorded and is available for Head of Line Blocking, Back Pressure and Agg. Flow Control Frames.

All 24 port ToR switches performed as expected— that is no HOL blocking was observed/measured during the L2 and L3 congestion test assuring that a congested port does not impact other ports, and thus the network, i.e., congestion is contained. In addition, the IBM BNT RackSwitch G8124 and Brocade VDX 6720-24 switches delivered 100% aggregated forwarding rate by implementing back pressure or signaling the Ixia test equipment with control frames to slow down the rate of packets entering the congested port, a normal and best practice for ToR switches. While not evident in the above graphic, uniquely the Arista 7124SX did indicate to Ixia test gear

© Lippis Enterprises, Inc. 2011

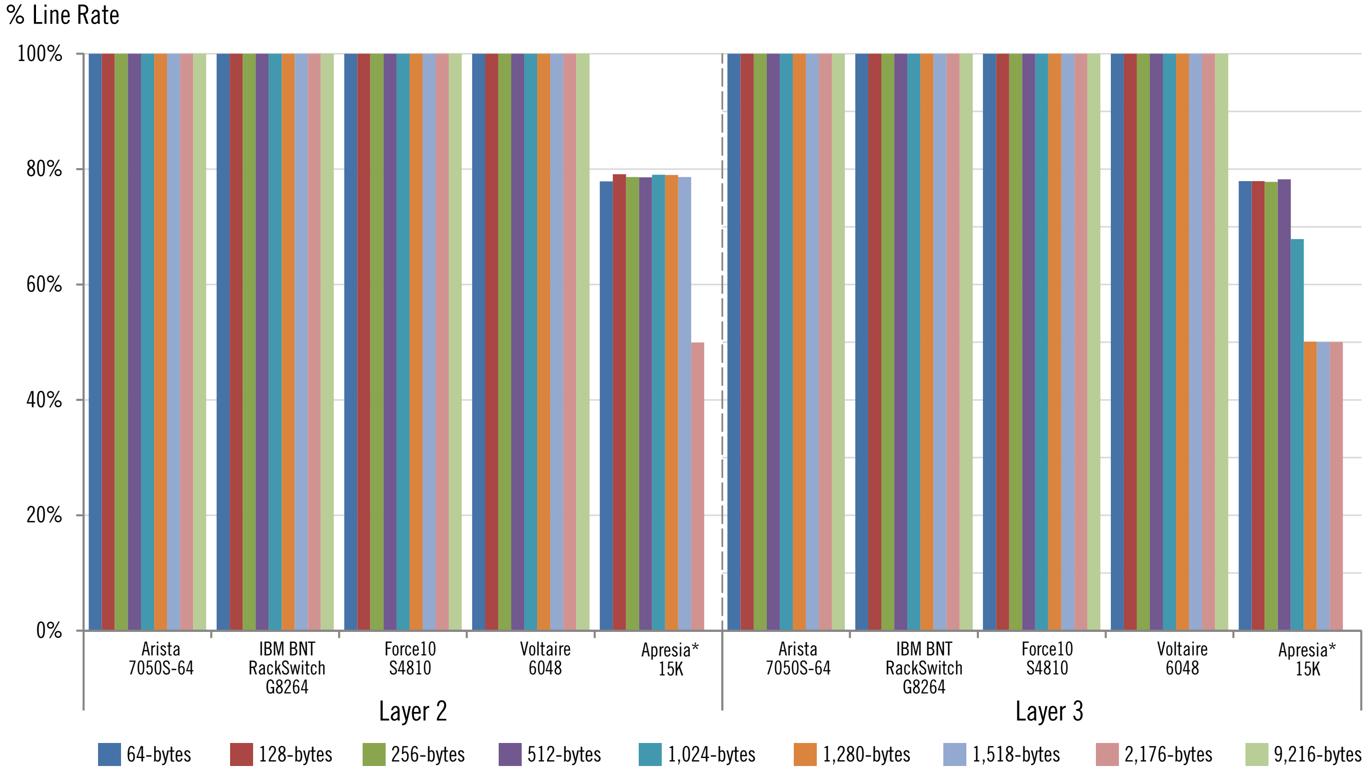
24 Port ToR Switches RFC 2889 Congestion Test
150% of line rate into single 10GbE

that it was using back pressure, however there were no flow control frames detected and in fact there were none. Ixia and other test equipment calculate back pressure per RFC 2889 paragraph 5.5.5.2., which states that if the total number of received frames on the congestion port surpasses the number of transmitted frames at MOL (Maximum Offered Load) rate then back pressure is present.

Thanks to the 7124SX's generous and dynamic buffer allocation it can overload ports with more packets than the MOL, therefore, the Ixia or any test equipment "calculates/sees" back pressure, but in reality this is an anomaly of the RFC testing method and not the 7124SX. The Arista 7124SX has a 2MB packet buffer pool and uses Dynamic Buffer Allocation (DBA) to manage congestion. Unlike other architectures that have a per-port fixed packet memory, the 7124SX can handle microbursts to allocate packet memory to the port that needs it. Such microbursts of traffic can be generated by RFC 2889 congestion tests when multiple traffic sources send traffic to the same destination for a short period of time. Using DBA, a port on the 7124SX can buffer up to 1.7MB of data in its transmit queue.

© Lippis Enterprises, Inc. 2011

ToR Switches RFC 2889 Congestion Test
150% of line rate into single 10GbE



	Layer 2															Layer 3														
Frames/ Packet Size (bytes)	Arista 7050S-64			IBM BNT RackSwitch G8264			Force10 S4810			Voltaire 6048			Apresia 15K			Arista 7050S-64			IBM BNT RackSwitch G8264			Force10 S4810			Voltaire 6048			Apresia 15K		
	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP
64	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	77.9%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	77.9%	N	Y
128	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	79.1%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	77.9%	N	Y
256	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	78.6%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	77.8%	N	Y
512	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	78.6%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	78.2%	N	Y
1,024	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	79.0%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	67.9%	Y	N
1,280	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	79.0%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	50.1%	Y	N
1,518	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	78.6%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	50.0%	Y	N
2,176	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	49.9%	Y	N	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	50.0%	Y	N
9,216	100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*			100%	N	Y	100%	N	Y	100%	N	Y	100%	N	Y	*		

* largest frame size supported on Apresia 15K is 9044. AFR% = Agg Forwarding Rate (% Line Rate) HOL = Head of Line Blocking BP = Back Pressure
Information was recorded and is available for Head of Line Blocking, Back Pressure and Agg. Flow Control Frames.

For the Arista 7050S-64, IBM BNT RackSwitch G8164, Force10 S4810 and Volaire Vantage™ 6048, these products performed as expected— that is no HOL blocking during the L2 and L3 congestion test assuring that a congested port does not impact other ports, and thus the network, i.e., congestion is contained. In addition, these switches delivered 100% aggregated forwarding rate by implementing back pressure or signaling the Ixia test equipment with control frames to slow down the rate of packets entering the congested port, a normal and best practice for ToR switches. While not evident in the above graphic, uniquely the Arista 7050S-64 did indicate to Ixia test gear

© Lippis Enterprises, Inc. 2011

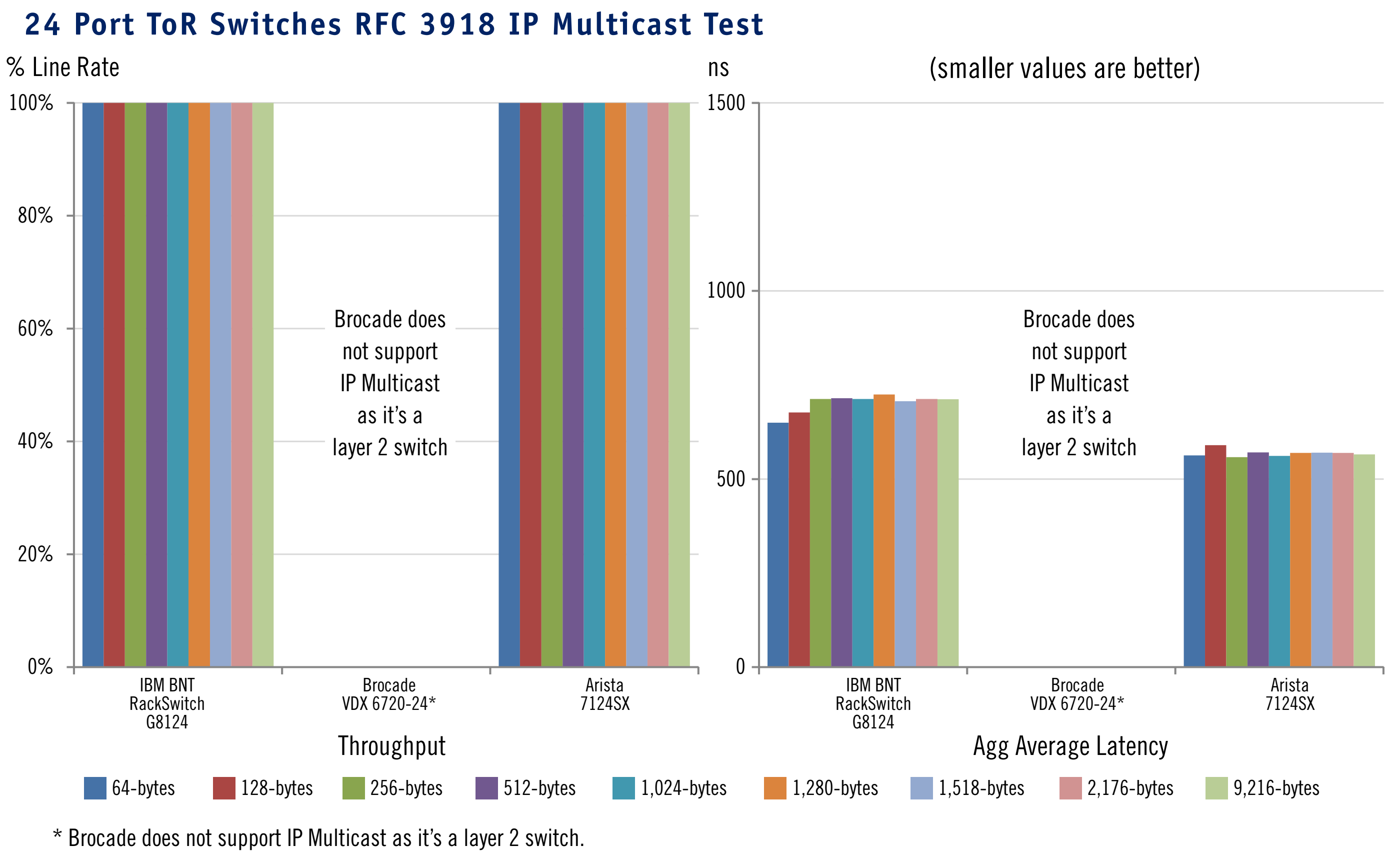
ToR Switches RFC 2889 Congestion Test
150% of line rate into single 10GbE

that it was using back pressure, however there were no flow control frames detected and in fact there were none. Ixia and other test equipment calculate back pressure per RFC 2889 paragraph 5.5.5.2., which states that if the total number of received frames on the congestion port surpasses the number of transmitted frames at MOL (Maximum Offered Load) rate then back pressure is present.

Thanks to the 7050S-64's generous and dynamic buffer allocation it can overload ports with more packets than the MOL, therefore, the Ixia or any test equipment "calculates/sees" back pressure, but in reality this is an anomaly of the RFC testing method and not the 7050S-64. The Arista 7050S-64 is designed with a dynamic buffer pool allocation such that during a microburst of traffic as during RFC 2889 congestion test when multiple traffic sources are destined to the same port, packets are buffered in packet memory. Unlike other architectures that have fixed per-port packet memory, the 7050S-64 uses Dynamic Buffer Allocation (DBA) to allocate packet memory to ports that need it. Under congestion, packets are buffered in shared packet memory of 9 MBytes. The 7050S-64 uses DBA to allocate up to 5MB of packet memory to a single port for lossless forwarding as observed during this RFC 2889 congestion test.

Apresia demonstrated HOL blocking at various packet sizes for both L3 and L2 forwarding. In addition, performance degradation beyond packet size 1518 during L2 congestion test and beyond packet size 512 during L3 congestion test were observed/detected.

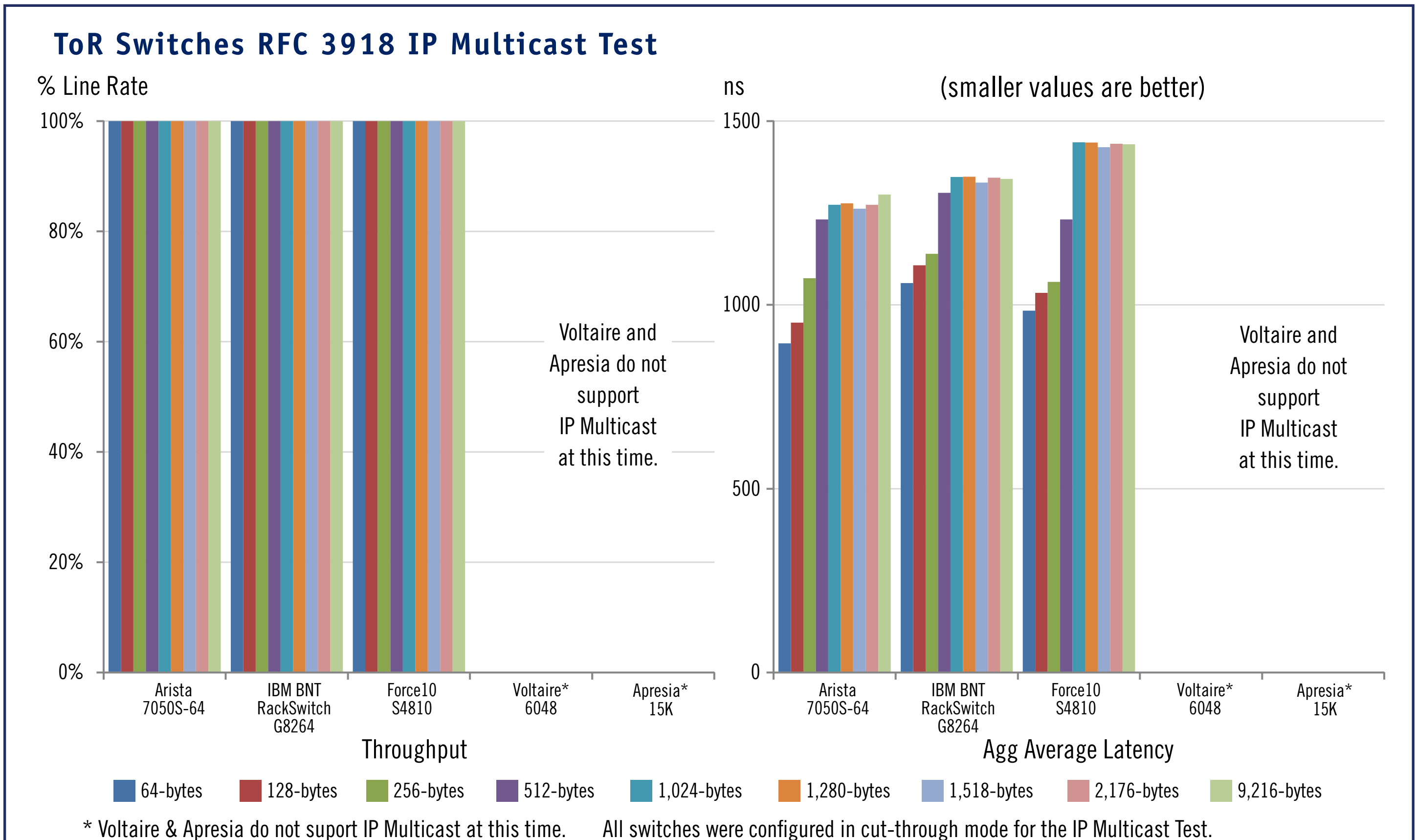
© Lippis Enterprises, Inc. 2011



	Throughput (% Line Rate)			Agg Average Latency (ns)		
Frames/Packet Size (bytes)	IBM BNT RackSwitch G8124	Brocade VDX 6720-24	Arista 7124SX	IBM BNT RackSwitch G8124	Brocade VDX 6720-24	Arista 7124SX
64	100%	Brocade does not support IP Multicast as it's a layer 2 switch	100%	650.00	Brocade does not support IP Multicast as it's a layer 2 switch	562.82
128	100%		100%	677.00		590.09
256	100%		100%	713.00		558.44
512	100%		100%	715.00		570.91
1,024	100%		100%	713.00		561.63
1,280	100%		100%	725.00		569.22
1,518	100%		100%	707.00		569.73
2,176	100%		100%	713.00		569.17
9,216	100%		100%	712.00		565.65

For the Arista 7124SX and IBM BNT RackSwitch G8124 these products performed as expected in RFC 3918 IP Multicast test delivering 100% throughput with zero packet loss. The Arista 7124SX demonstrated over 100ns less latency than the IBM BNT RackSwitch G8124 at various packet sizes. Notice that both the Arista 7124SX and IBM BNT RackSwitch G8124 demonstrated consistently low latency across all packet sizes. The Brocade VDX™ 6720-24 is a layer 2 cut through switch and thus did not participate in the IP Multicast test.

© Lippis Enterprises, Inc. 2011

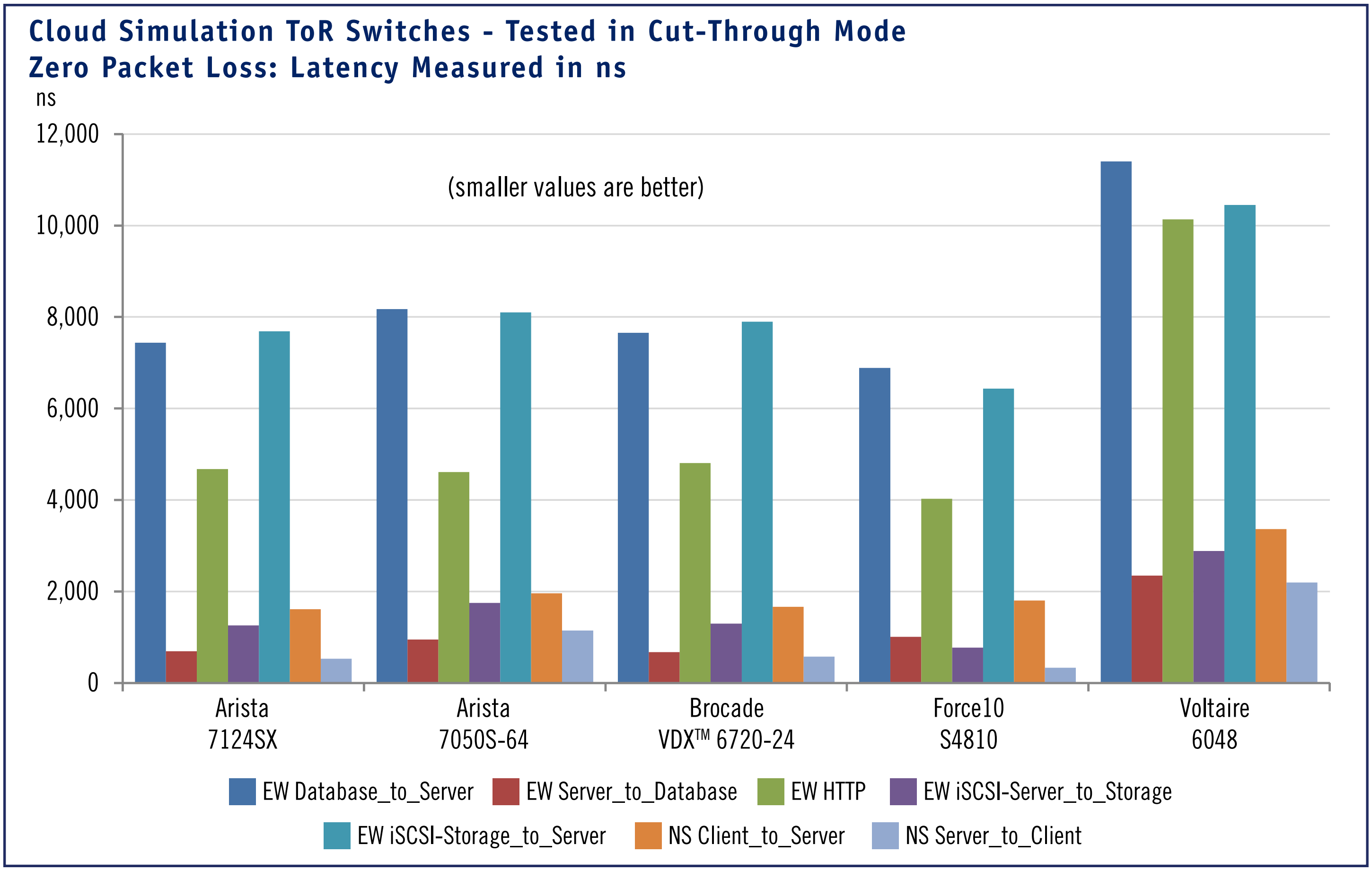


	Throughput (% Line Rate)					Agg Average Latency (ns)				
Frames/Packet Size (bytes)	Arista 7050S-64	IBM BNT RackSwitch G8264	Force10 S4810	Voltaire 6048	Apresia 15K	Arista 7050S-64	IBM BNT RackSwitch G8264	Force10 S4810	Voltaire 6048	Apresia 15K
64	100%	100%	100%	Voltaire and Apresia do not support IP Multicast at this time.		894.90	1059.00	984.00	Voltaire and Apresia do not support IP Multicast at this time.	
128	100%	100%	100%			951.95	1108.00	1033.00		
256	100%	100%	100%			1072.35	1139.00	1062.00		
512	100%	100%	100%			1232.06	1305.00	1232.00		
1,024	100%	100%	100%			1272.03	1348.00	1442.00		
1,280	100%	100%	100%			1276.35	1348.00	1442.00		
1,518	100%	100%	100%			1261.49	1332.00	1429.00		
2,176	100%	100%	100%			1272.46	1346.00	1438.00		
9,216	100%	100%	100%			1300.35	1342.00	1437.00		

All switches were configured in cut-through mode for the IP Multicast Test.

For the Arista 7050S-64, IBM BNT RackSwitch G8264 and Force10 S4810, these products performed as expected in RFC 3918 IP Multicast test delivering 100% throughput with zero packet loss. In terms of IP Multicast "average latency" between the Arista 7050S-64, IBM BNT RackSwitch G8264 and Force10 S4810 their data is similar with higher latencies at larger packet sizes. There is a slight difference in aggregate average latency between the Arista 7050S-64 and IBM BNT RackSwitch G8264 with Arista offering 100 ns less latency with the Force10 S4810 offering higher IP Multicast in aggregate average latency at the larger packet sizes. These three ToR switches can be configured in 48 10GbE ports plus 2 40GbE uplink ports or 64 10GbE ports. Voltaire and Apresia does not support IP Multicast at this time.

© Lippis Enterprises, Inc. 2011

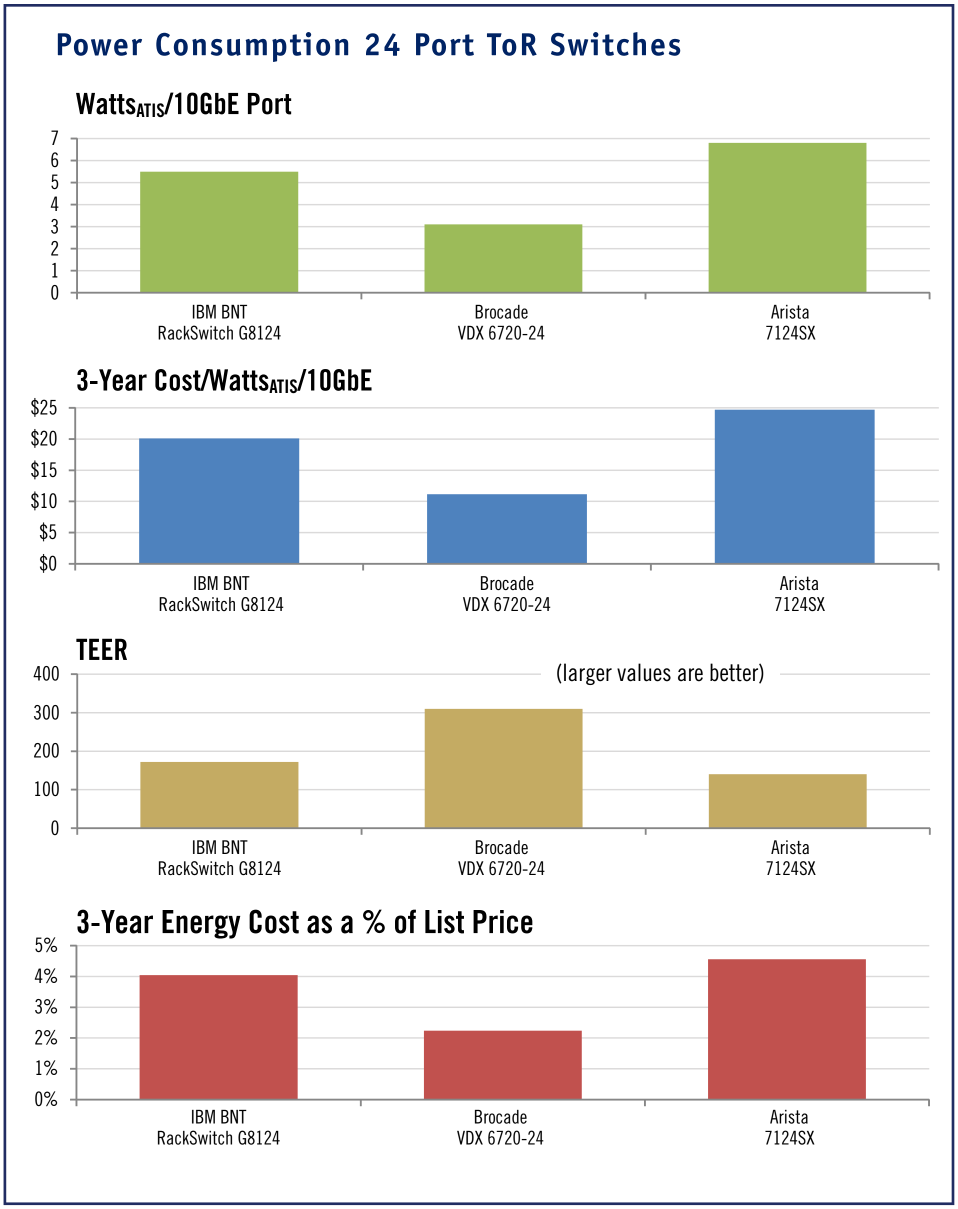


Cloud Simulation ToR Switches
Zero Packet Loss: Latency Measured in ns

Company	Product	EW Database_ to_Server	EW Server_ to_Database	EW HTTP	EW iSCSI- Server_to_ Storage	EW iSCSI- Storage_to_ Server	NS Client_ to_Server	NS Server_ to_Client
Arista	7124SX	7439	692	4679	1256	7685	1613	530
Arista*	7050S-64	8173	948	4609	1750	8101	1956	1143
Brocade	VDX™ 6720-24	7653	675	4805	1298	7901	1666	572
Force10*	S4810	6890	1010	4027	771	6435	1801	334
Voltaire	6048	11401	2348	10134	2884	10454	3366	2197

The Arista 7124SX, Arista 7050S-64, Brocade VDX™ 6720-24, Force10 S4810 and Voltaire Vantage™ 6048 delivered 100% throughput with zero packet loss and nanosecond latency during the cloud simulation test. The Arista 7124SX, Arista 7050S-64, Brocade VDX™ 6720-24 and Force10 S4810 demonstrated similar latencies across the various cloud protocols. All demonstrated spikes at EW Database-to-Server, EW HTTP and EW iSCSI-Storage-to-Server cloud protocols. The Voltaire Vantage™ 6048 added between 600ns and 300ns to Arista and Force10's latency results.

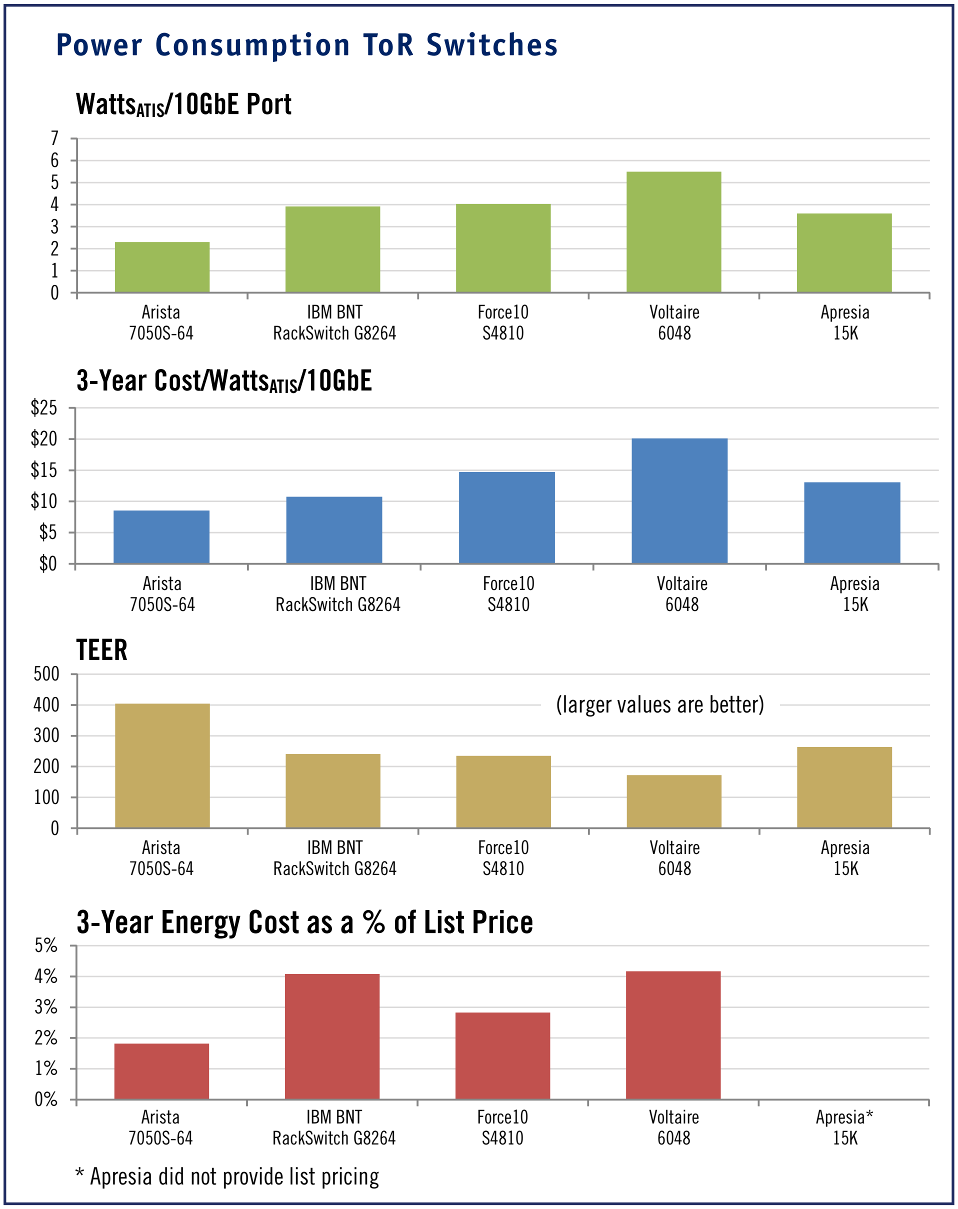
© Lippis Enterprises, Inc. 2011



The Brocade VDX™ 6720-24 is the standout in this group of ToR switches with a low 3.1 $WATTS_{ATIS}$ and a TEER value of 310. The Arista 7124SX and IBM BNT RackSwitch G8124 did not differ materially in their power consumption and 3-Year energy cost as a percent of list pricing. All 24 port ToR switches are highly energy efficient with high TEER values and low $WATTS_{ATIS}$.

Company	Product	$Watts_{ATIS}$/port	Cost/ $Watts_{ATIS}$/Year	3 yr Energy Cost per 10GbE Port	TEER	3 yr Energy Cost as a % of List Price	Cooling
IBM	BNT Rack Switch G8124	5.5	$6.70	$20.11	172	4.04%	Front-to-Back
Brocade	VDX™ 6720-24	3.1	$3.72	$11.15	310	2.24%	Front-to-Back, Back-to-Front
Arista	7124SX	6.8	$8.24	$24.71	140	4.56%	Front-to-Back, Back-to-Front

© Lippis Enterprises, Inc. 2011



The Arista 7050S-64 is the standout in this group of ToR switches with a low 2.3 $WATTS_{ATIS}$ and a TEER value of 404, both of which were the lowest and highest we have ever measured making the Arista 7050S-64 the most power efficient switch measured to date. All switches are highly energy efficient switches with high TEER values and low $WATTS_{ATIS}$. While there is difference in the three-year energy cost as a percentage of list price, this is due to the fact these products were configured differently during the Lippis/Ixia test with the Arista 7050S-64 and IBM BNT RackSwitch G8264 configured as 64 10GbE while the Force10 S4810 configured as 48 10GbE ports with subsequent different list pricing. Hitachi Cable was unable to provide list pricing information thus there is no data on the Apresia 15K's power cost per 10GbE, three-year energy cost as a percentage of list price.

Company	Product	$Watts_{ATIS}$/port	Cost/ $Watts_{ATIS}$/Year	3 yr Energy Cost per 10GbE Port	TEER	3 yr Energy Cost as a % of List Price	Cooling
Arista	7050S-64	2.3	$2.85	$8.56	404	1.82%	Front-to-Back, Back-to-Front
IBM	BNT Rack Switch G8264	3.9	$4.78	$10.75	241	4.08%	Front-to-Back
Force10	S4810	4.0	$4.91	$14.73	235	2.83%	Front-to-Back
Voltaire	6048	5.5	$6.70	$20.11	172	4.17%	Front-to-Back
Apresia*	15K	3.6	$13.09	$13.09	264		Front-to-Back

* Apresia did not provide list pricing

© Lippis Enterprises, Inc. 2011

ToR Industry Recommendations

The following provides a set of recommendations to IT business leaders and network architects for their consideration as they seek to design and build their private/public data center cloud network fabric. Most of the recommendations are based upon our observations and analysis of the test data. For a few recommendations, we extrapolate from this baseline of test data to incorporate key trends and how these ToR switches may be used in their support for corporate advantage.

10GbE ToR Switches Ready for Deployment: 10GbE ToR switches are ready for prime time, delivering full line rate throughput at zero packet loss and nanosecond latency plus single to double-digit delay variation. We are now in the era of 500ns latency ToR switching. In addition, these ToR switches offer low power consumption with energy cost over a three-year period estimated between 1.8% and 4% of acquisition cost.

Evaluate Each ToR Switch Separately: While this set of ToR switch category tested very well, there are differences between vendors especially in the congestion and cloud simulation test. Therefore, it is recommended to review each supplier's results and make purchase decisions accordingly.

Deploy ToR Switches that Demonstrate Efficient Power and Cooling: In addition to high TEER values and low $WATTS_{ATIS}$ all ToR switches tested support front-to-back or rear-to-front cooling in support of data center hot/cold aisle airflow designs. In fact some offer a reversible option too. Therefore, it is recommended that these ToR switches can be used as part of an overall green data center initiative.

Ready for Storage Enablement: Most of the ToR switches demonstrated the performance and latency required to support storage enablement or converged I/O. In fact, all suppliers have invested in storage enablement, such as support for CEE, FCoE, iSCSI, ATA, NAS, etc., and while these features were not tested in the Lippis/Ixia evaluation, these switches demonstrated that the raw capacity is built into the switches for its support.

Evaluate ToR Switches for Network Services and Fabric Differences: There are differences between suppliers in terms of the network services they offer as well as how their ToR switches connect to Core/Spine switches to create a data center fabric. It is recommended that IT business leaders evaluate ToR switches with Core switches to assure that the network services and fabric attributes sought after are realized.

Single Chip Design Advantages: With most, but not all ToR switches being designed with merchant silicon and proving their performance and power consumption advantages, expect many other suppliers to introduce ToR switches based upon this design. Competitive differentiation will quickly shift toward network operating system and network services support especially virtualization aware and unified networking.

Connect Servers at 10GbE: The Lippis/Ixia test demonstrated the performance and power consumption advantages of 10GbE networking, which can be put to work and exploited for corporate advantage. For new server deployments in private/public data center cloud networks, 10GbE is recommended as the primary network connectivity service as a network fabric exists to take full advantage of server I/O at 10GbE bandwidth and latency levels.

40GbE Uplink Options Become Available: The Arista 7050S-64, IBM BNT RackSwitch G8264 and Force10 S4810 support four 40GbE ports. Two 40GbE configuration options are supported: 1) four 40GbE uplinks or 2) breakout each 40GbE port to four 10GbE ports to drive up 10GbE port density to 64. The Lippis/Ixia evaluation tested the Arista 7050S-64 and IBM BNT RackSwitch G8264 using its 40GbE as a fan-in of 4 10GbE to increase 10GbE port density to 64. This option offers design flexibility at no cost to performance, latency or power consumption, and thus we recommend its use. Further, in the fall of 2011 we expect many new 40GbE core switches to be available to accept these 40GbE uplinks.

© Lippis Enterprises, Inc. 2011

Core/Spine Cross-Vendor Analysis

There were three Core/Spine Switches evaluated for performance and power consumption in the Lippis/Ixia test. These participating vendors were:

Alcatel-Lucent OmniSwitch 10K

Arista 7504 Series Data Center Switch

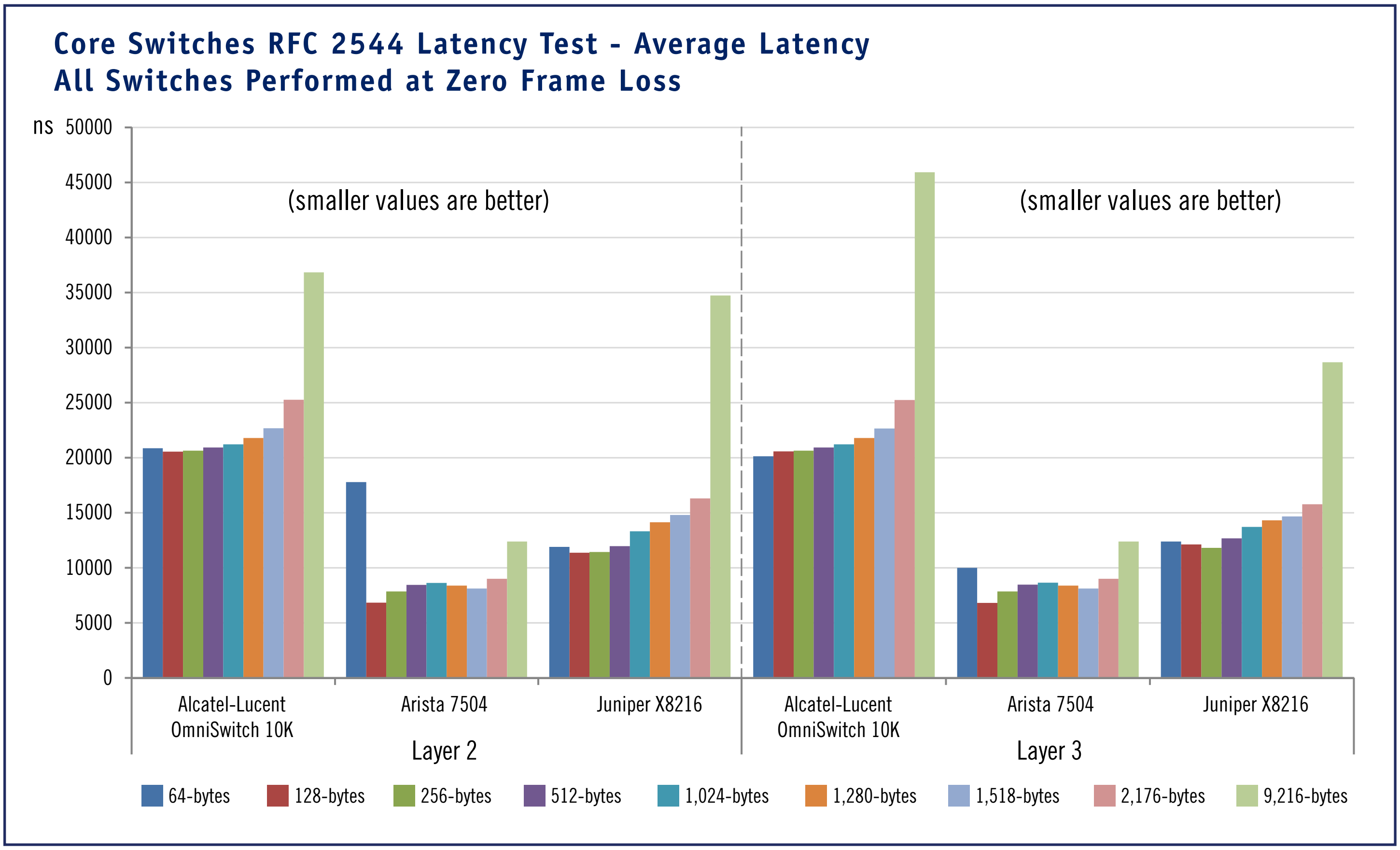
Juniper Network EX Series EX8200 Ethernet Switch

These switches represent the state-of-the-art of computer network hardware and software engineering, and are central to private/public data center cloud computing infrastructure. If not for this category of Ethernet switching, cloud computing would not exist. The Lippis/Ixia public test was the first evaluation for every Core switch tested. Each supplier's Core switch was evaluated for its fundamental performance and power consumption features. The Lippis/Ixia test results demonstrate that these new Core switches provide state-of-the-art performance at efficient power consumption levels not seen before. The port density tested for these Core switches ranged from 128 10GbE ports to a high of 256 10GbE.

IT business leaders are responding favorably to Core switches equipped with a value proposition of high performance, high port density, competitive acquisition cost, virtualization aware services, high reliability and low power consumption. These Core switches currently are in high demand with quarterly revenues for mid-size firms in the $20 to $40M plus range. The combined market run rate for both ToR and Core 10GbE switching is measured in the multi-billion dollar range. Further, Core switch price points on a 10GbE per port basis are a low of $1,200 to a high of $6,093. Their list price varies from $230K to $780K with an average order usually being in the million plus dollar range. While there is a large difference in list price as well as price per port between vendors, the reason is found in the number of network services supported by the various suppliers and 10GbE port density.

We compare each of the above firms in terms of their ability to forward packets: quickly (i.e., latency), without loss or their throughput at full line rate, when ports are oversubscribed with network traffic by 150%, in IP multicast mode and in cloud simulation. We also measure their power consumption as described in the Lippis Test Methodology section above.

© Lippis Enterprises, Inc. 2011



	Layer 2			Layer 3		
Frames/ Packet Size (bytes)	Alcatel-Lucent* OmniSwitch 10K	Arista 7504	Juniper EX8126	Alcatel-Lucent* OmniSwitch 10K	Arista 7504	Juniper EX8126
64	20864.00	17791.20	11891.60	20128.00	9994.88	12390.38
128	20561.00	6831.50	11366.20	20567.00	6821.32	12112.50
256	20631.00	7850.70	11439.60	20638.00	7858.69	11814.05
512	20936.00	8452.00	11966.80	20931.00	8466.53	12685.41
1,024	21216.00	8635.50	13322.90	21211.00	8647.96	13713.25
1,280	21787.00	8387.00	14131.80	21793.00	8387.26	14318.48
1,518	22668.00	8122.10	14789.60	22658.00	8112.95	14664.84
2,176	25255.00	8996.00	16302.70	25242.00	8996.55	15767.44
9,216	36823.00	12392.80	34722.70	45933.00	12387.99	28674.08

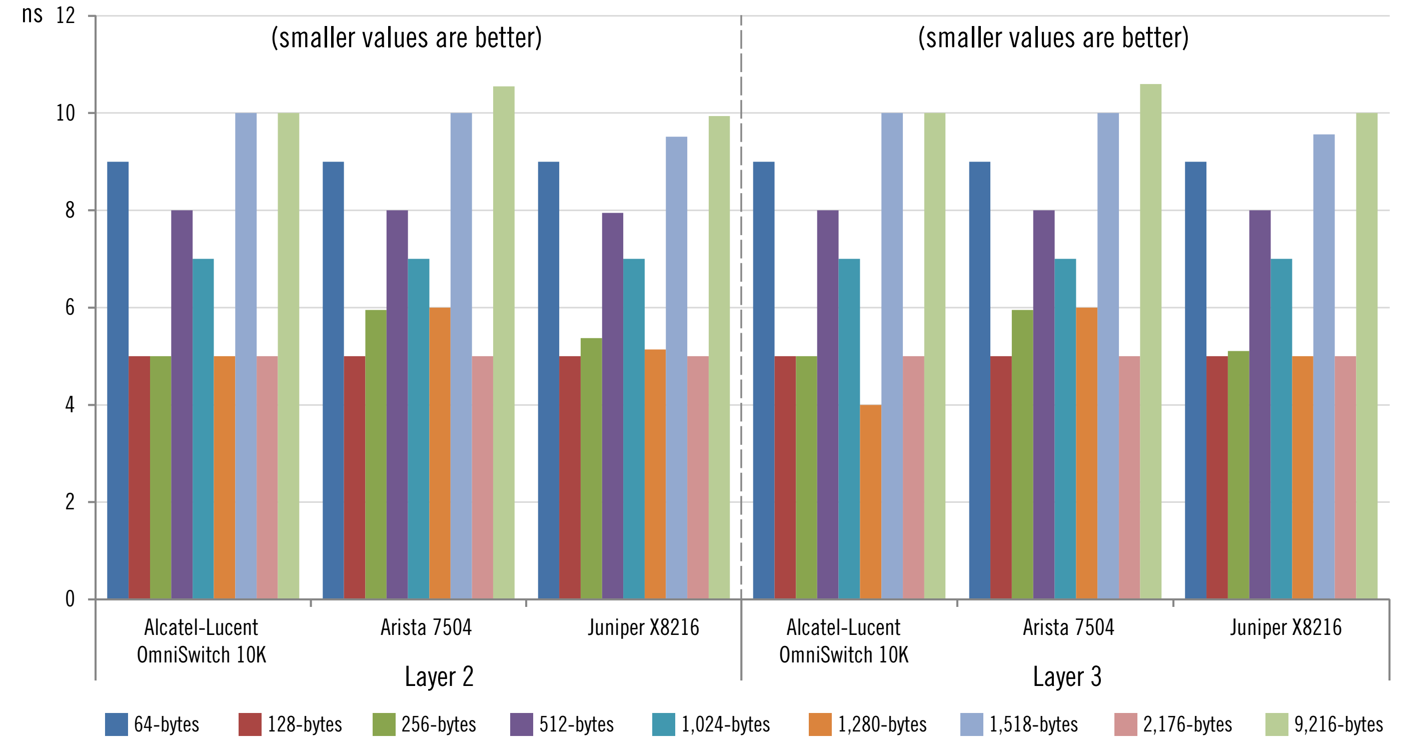
* RFC2544 was conducted on the OmniSwitch 10K via default configuration, however Alcatel-Lucent has an optimized configuration for low latency networks which it says improves these results.

We show average latency and average delay variation across all packet sizes for layer 2 and 3 forwarding. Measurements were taken from ports that were far away from each other to demonstrate the consistency of performance between modules and across their backplanes.

One standout here is the Arista 7504's large latency at 64Byte size packets. According to Arista, the Arista 7500 series performance and design are fine tuned for applications its customers use. These applications use mixed packet sizes, rather than an artificial stream of wire speed 64 byte packets.

© Lippis Enterprises, Inc. 2011

Core Switches RFC 2544 Latency Test - Average Delay Variation
All Switches Performed at Zero Frame Loss

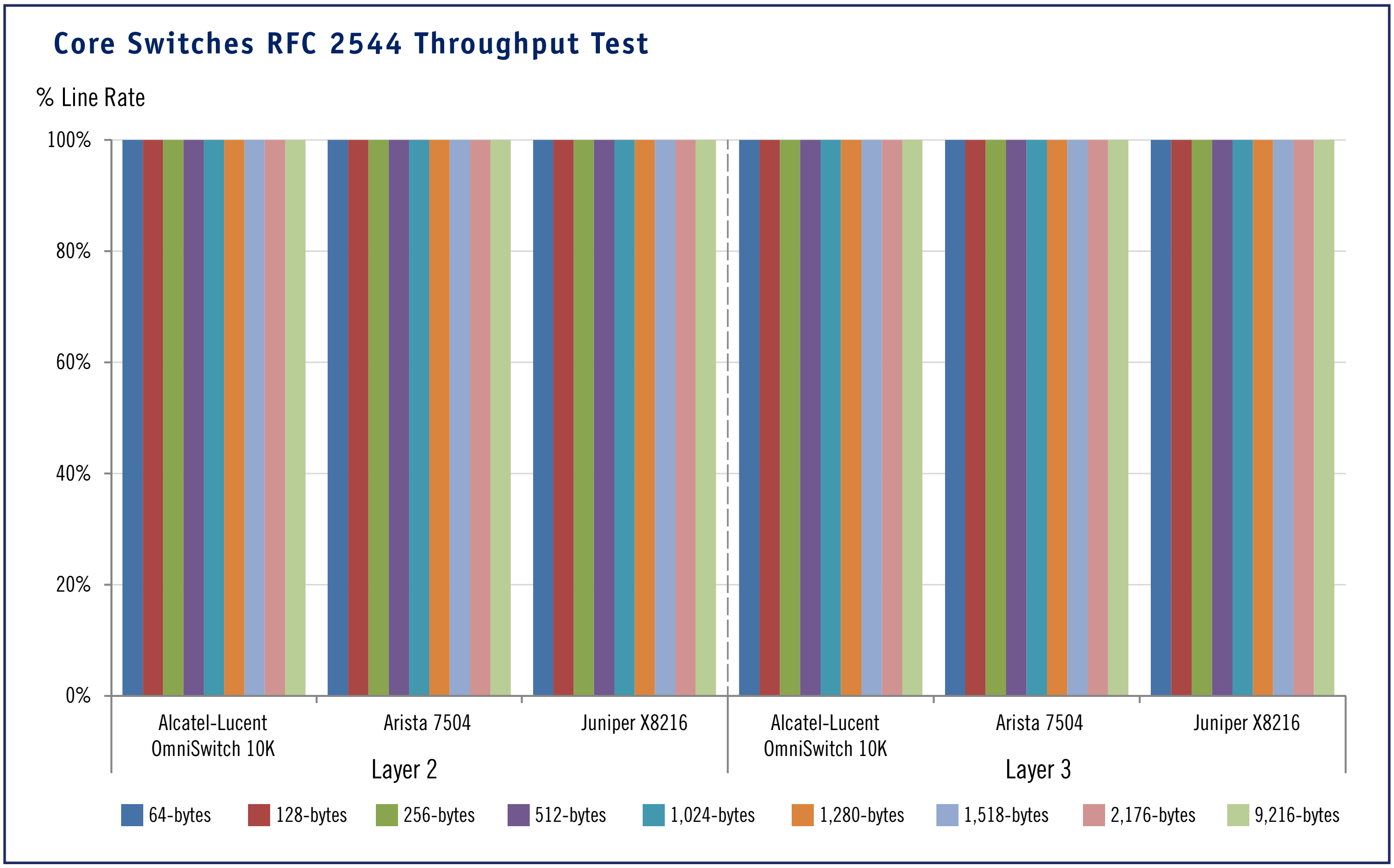


	Layer 2			Layer 3		
Frames/ Packet Size (bytes)	Alcatel-Lucent* OmniSwitch 10K	Arista 7504	Juniper EX8126	Alcatel-Lucent* OmniSwitch 10K	Arista 7504	Juniper EX8126
64	9.00	9.00	9.00	9.00	9.00	9.00
128	5.00	5.00	5.00	5.00	5.00	5.00
256	5.00	6.00	5.40	5.00	5.95	5.11
512	8.00	8.00	7.90	8.00	8.00	8.00
1,024	7.00	7.00	7.00	7.00	7.00	7.00
1,280	5.00	6.00	5.10	4.00	6.00	5.00
1,518	10.00	10.00	9.50	10.00	10.00	9.56
2,176	5.00	5.00	5.00	5.00	5.00	5.00
9,216	10.00	10.50	9.90	10.00	10.60	10.00

* RFC2544 was conducted on the OmniSwitch 10K via default configuration, however Alcatel-Lucent has an optimized configuration for low latency networks which it says improves these results.

Notice there is little difference in average delay variation across all vendors proving consistent latency under heavy load at zero packet loss for L2 and L3 forwarding. Average delay variation is in the 5 to 10 ns range. Difference does reside within average latency between suppliers, thanks to different approaches to buffer management and port densities.

© Lippis Enterprises, Inc. 2011

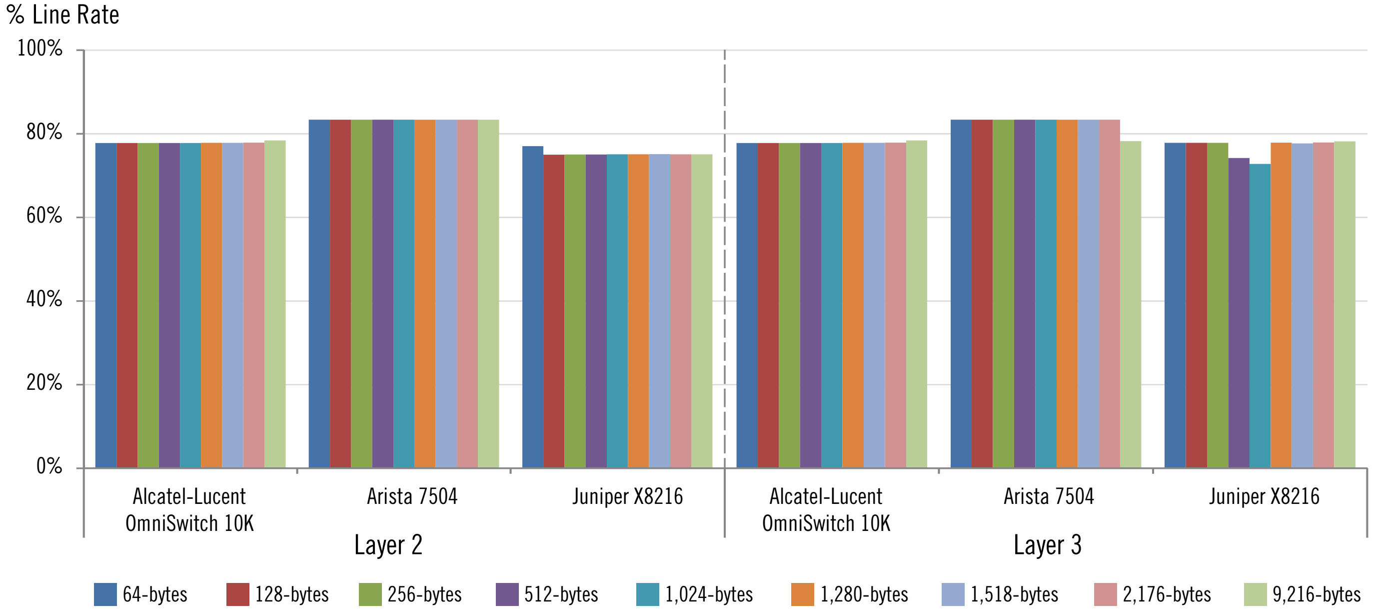


	Layer 2			Layer 3		
Frames/ Packet Size (bytes)	Alcatel-Lucent OmniSwitch 10K	Arista 7504	Juniper EX8126	Alcatel-Lucent OmniSwitch 10K	Arista 7504	Juniper EX8126
64	100%	100%	100%	100%	100%	100%
128	100%	100%	100%	100%	100%	100%
256	100%	100%	100%	100%	100%	100%
512	100%	100%	100%	100%	100%	100%
1,024	100%	100%	100%	100%	100%	100%
1,280	100%	100%	100%	100%	100%	100%
1,518	100%	100%	100%	100%	100%	100%
2,176	100%	100%	100%	100%	100%	100%
9,216	100%	100%	100%	100%	100%	100%

As expected, all switches are able to forward L3 packets at all sizes at 100% of line rate with zero packet loss, proving that these switches are high performance devices.

© Lippis Enterprises, Inc. 2011

Core Switches RFC 2889 Layer 2 Congestion Test
150% of Line Rate into Single 10GbE



	Layer 2									Layer 3								
Frames/ Packet Size (bytes)	Alcatel-Lucent OmniSwitch 10K			Arista 7504			Juniper EX8126			Alcatel-Lucent OmniSwitch 10K			Arista 7504			Juniper EX8126		
	AFR%	HOL	BP	AFR%	HOL	BP*	AFR%	HOL	BP	AFR%	HOL	BP	AFR%	HOL	BP*	AFR%	HOL	BP
64	77.80%	N	N	83.34%	N	Y	77.01%	N	N	77.80%	N	N	84.34%	N	Y	77.81%	N	N
128	77.80%	N	N	83.34%	N	Y	75.01%	N	N	77.80%	N	N	84.34%	N	Y	77.84%	N	N
256	77.80%	N	N	83.35%	N	Y	75.02%	N	N	77.80%	N	N	84.35%	N	Y	77.83%	N	N
512	77.80%	N	N	83.35%	N	Y	75.03%	N	N	77.80%	N	N	84.35%	N	Y	74.20%	N	N
1,024	77.80%	N	N	83.36%	N	Y	76.06%	N	N	77.80%	N	N	84.36%	N	Y	72.78%	N	N
1,280	77.80%	N	N	83.36%	N	Y	76.06%	N	N	77.80%	N	N	84.36%	N	Y	77.87%	N	N
1,518	77.80%	N	N	83.36%	N	Y	75.12%	N	N	77.80%	N	N	84.36%	N	Y	77.69%	N	N
2,176	77.90%	N	N	83.36%	N	Y	76.06%	N	N	77.90%	N	N	84.36%	N	Y	77.91%	N	N
9,216	78.40%	N	N	83.36%	N	Y	75.06%	N	N	78.40%	N	N	78.21%	Y	N	78.16%	N	N

AFR% = Agg Forwarding Rate (% Line Rate) HOL = Head of Line Blocking BP = Back Pressure

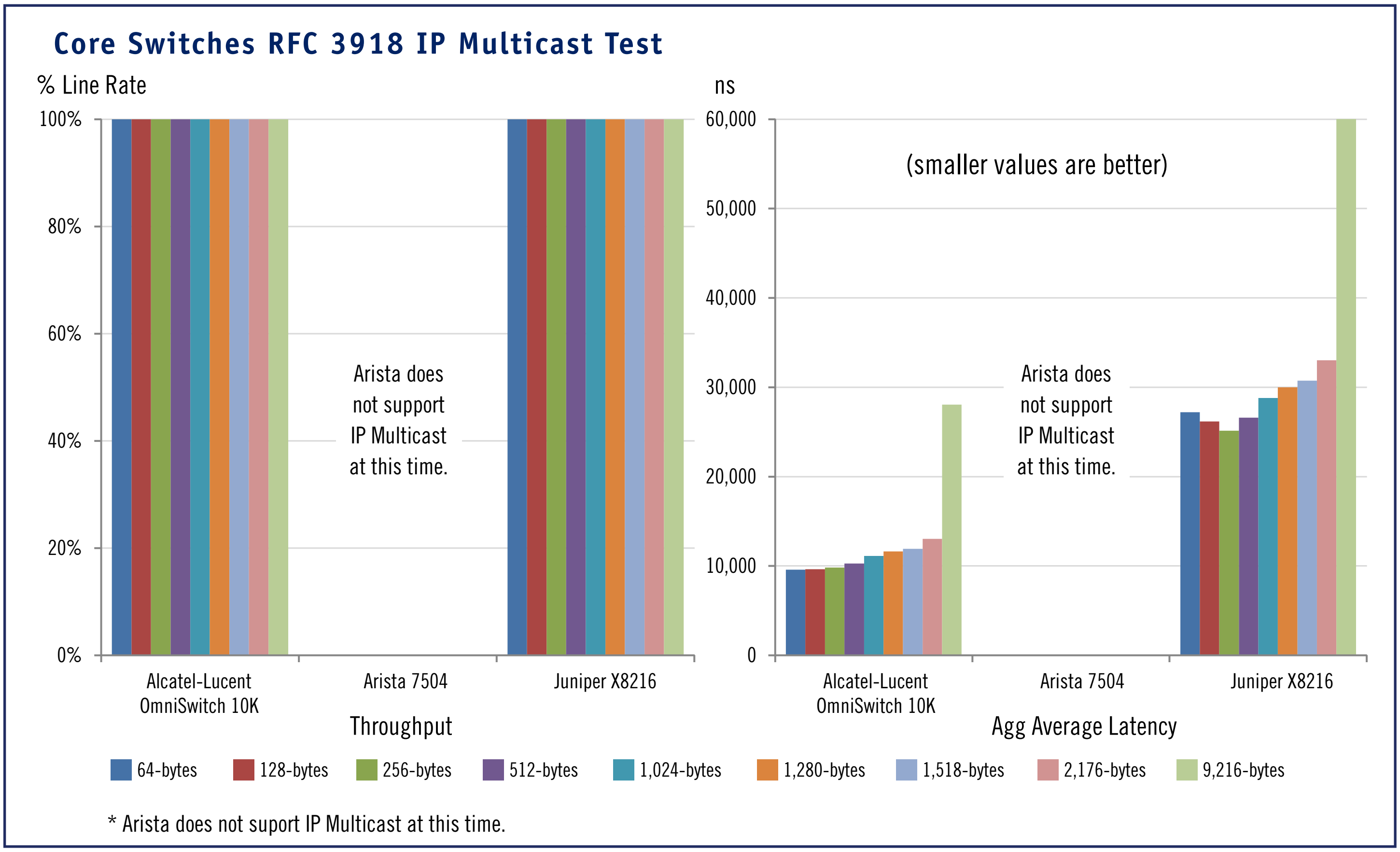
* Note that while the Arista's 7504 shows back pressure, in fact there is none. All test equipment including IXIA calculates back pressure per RFC 2889 paragraph 5.5.5.2. which states that if the total number of received frames on the congestion port surpasses the number of transmitted frames at MOL (Maximum Offered Load) rate then back pressure is present. Thanks to the 7504's 2.3GB of packet buffer memory it can overload ports with more packets than the MOL, therefore, IXIA or any test equipment "calculates/sees" back pressure, but in reality this is an anomaly of the RFC testing method and not the 7504. The Arista 7504 can buffer up 40ms of traffic per port at 10GbE speeds which is 400K bits or 5,425 packets of 9216 bytes.

All Core switches performed extremely well under congestion conditions with nearly no HOL blocking and between 77 to 83% aggregated forwarding as a percentage of line rate. It's expected that at 150% of offered load to a port, that that Core switch's port would show 33% loss if it's receiving at 150% line rate and not performing back pressure.

A few standouts arose in this test. First, Arista's 7504 delivered nearly 84% of aggregated forwarding rate as percentage of line rate during congestion conditions for both L2 and L3 traffic flows, the highest for all suppliers. This is due to its generous 2.3GB of buffer memory design as well as its VOQ architecture. Note also that Arista is the only Core switch between Juniper and Alcatel-Lucent that showed HOL blocking for 9216 byte size packets at L3. Note the 7504 was in beta testing at the time of the Lippis/IXIA test and there was a "*corner case*" at 9216 bytes at L3 that needed further tuning. Arista states that its production code provides wire-speed L2/L3 performance at all packet sizes without any head of line blocking

Second, Alcatel-Lucent and Juniper delivered consistent and perfect performance with no HOL blocking or back pressure for all packet sizes during both L2 and L3 forwarding.

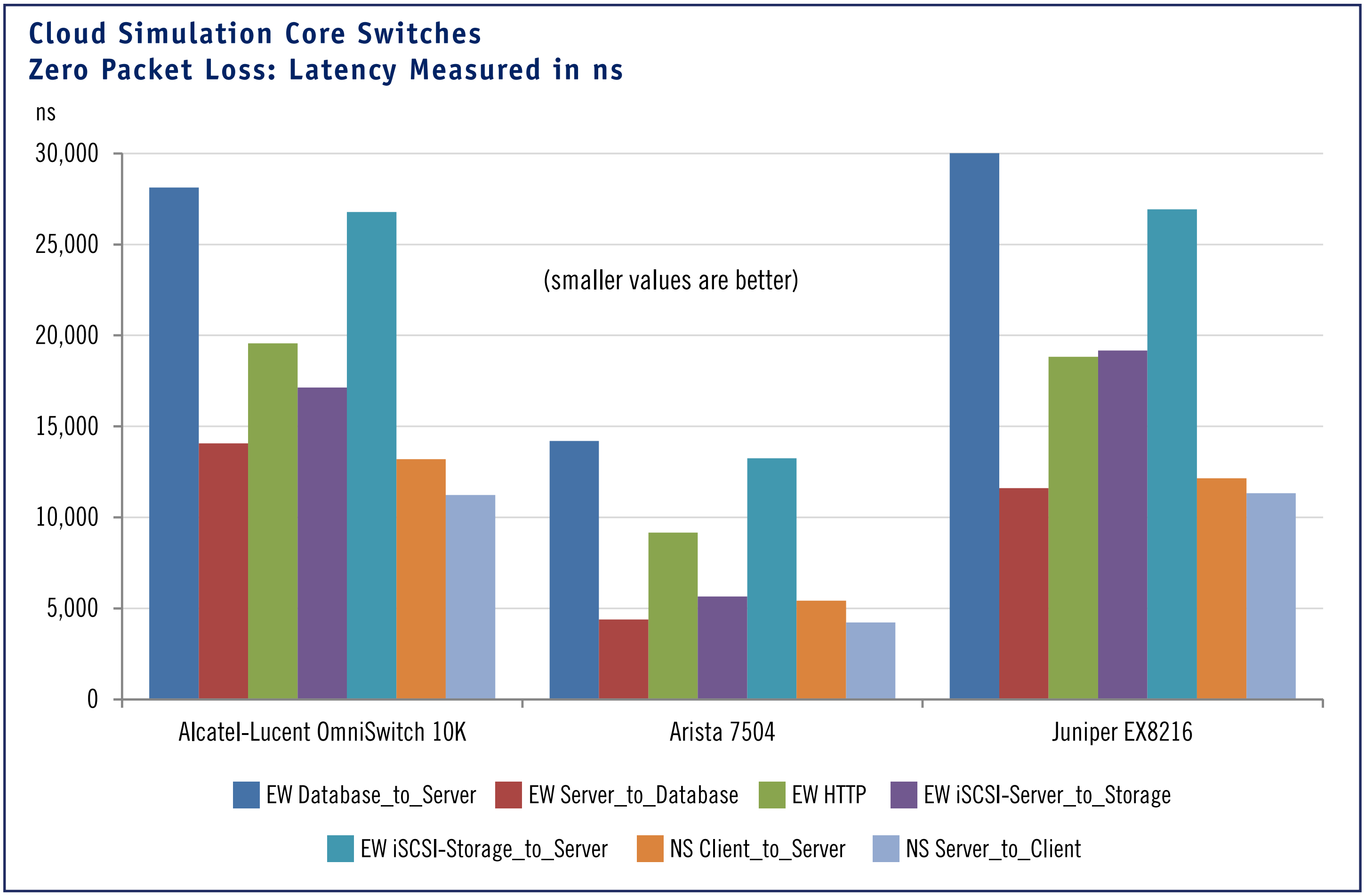
© Lippis Enterprises, Inc. 2011



	Throughput (% Line Rate)			Agg Average Latency (ns)		
Frames/ Packet Size (bytes)	Alcatel-Lucent OmniSwitch 10K	Arista 7504*	Juniper EX8126	Alcatel-Lucent OmniSwitch 10K	Arista 7504*	Juniper EX8126
64	100%		100%	9596		27210
128	100%		100%	9646		26172
256	100%		100%	9829		25135
512	100%	Arista does not support IP Multicast at this time.	100%	10264	Arista does not support IP Multicast at this time.	26600
1,024	100%		100%	11114		28802
1,280	100%		100%	11625		29986
1,518	100%		100%	11921		30746
2,176	100%		100%	13023		33027
9,216	100%		100%	28059		60055

Arista does not support IP Multicast at this time and thus is excluded from this test. Alcatel-Lucent's OmniSwitch 10K demonstrated 100% throughput at line rate and the shortest aggregated average latency. Juniper's EX8216 demonstrated 100% throughput at line rate and the IP multicast aggregated average latency twice that of Alcatel-Lucent's OmniSwitch.

© Lippis Enterprises, Inc. 2011

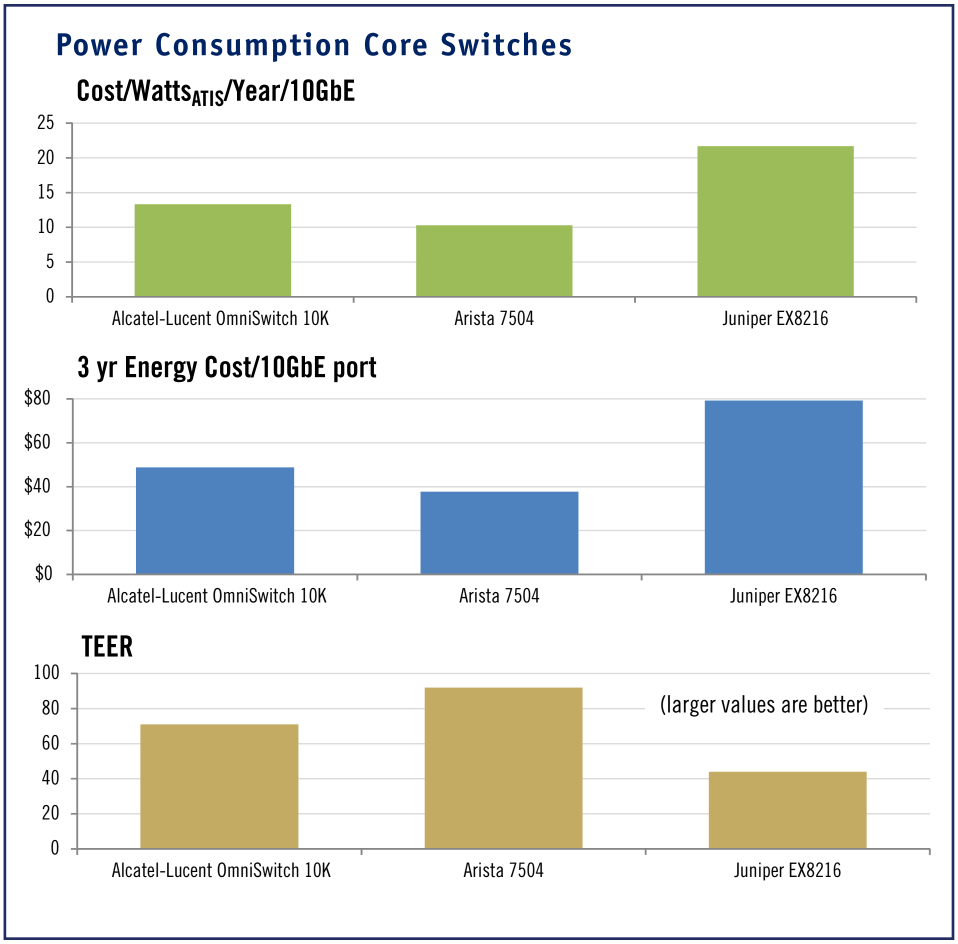


Cloud Simulation Core Switches
Zero Packet Loss: Latency Measured in ns

Company	Product	EW Database_ to_Server	EW Server_to_ Database	EW HTTP	EW iSCSI-Server_to_ Storage	EW iSCSI-Storage_to_ Server	NS Client_ to_Server	NS Server_ to_Client
Alcatel-Lucent	OmniSwitch 10K	28125	14063	19564	17140	26780	13194	11225
Arista	7504	14208	4394	9171	5656	13245	5428	4225
Juniper	EX8216	30263	11614	18833	19177	26927	12145	11322

All Core switches delivered 100% throughput with zero packet loss during the cloud simulation test. Alcatel-Lucent's and Juniper's average latencies for various traffic flows were consistent with each other while Arista's latency was measured, at specific traffic types, significantly lower.

© Lippis Enterprises, Inc. 2011



There are differences in the power efficiency across these Core switches. The OmniSwitch 10K, 7540 and EX8216 represent a breakthrough in power efficiency for Core switches with previous generation of switches consuming as much as 70 $WATTS_{ATIS}$. These switches consume between 10 and 21 $WATTS_{ATIS}$ with TEER values between 92 and 71; remember the higher the TEER value the better. The three year cost per 10GbE port shows the difference between the three core switches with 3-yr estimated cost per 10GbE ranging from a low of $37.69 to a high of $79.26. In terms of cooling, Juniper's EX8216 is the only Core switch that does not support front-to-back airflow. However, third party cabinet options from vendors such as Chatsworth support hot-aisle cold-aisle deployments of up to two EX8216 chassis in a single cabinet.

Company	Product	$Watts_{ATIS}$/ port	3-Yr Energy Cost per 10GbE Port	TEER	Cooling
Alcatel-Lucent	OmniSwitch 10K	13.3	$48.77	71	Front-to-Back
Arista	7504	10.3	$37.69	92	Front-to-Back
Juniper	EX8216	21.7	$79.26	44	Side-to-Side*

* 3rd party cabinet options from vendors such as Chatsworth support hot-aisle cold-aisle deployments of up to two EX8216 chassis in a single cabinet.

© Lippis Enterprises, Inc. 2011

Core Switch Industry Recommendations

The following provides a set of recommendations to IT business leaders and network architects for their consideration as they seek to design and build their private/public data center cloud network fabric. Most of the recommendations are based upon our observations and analysis of the test data. For a few recommendations, we extrapolate from this baseline of test data to incorporate key trends and how these Core switches may be used in their support for corporate advantage.

10GbE Core Switches Ready for Deployment: 10GbE Core switches are ready for prime time, delivering full line rate throughput at zero packet loss and nanosecond latencies plus single to double-digit delay variation. In addition, these Core switches offer low power consumption with energy cost over a three-year period estimated between 1.3% and 3.14% of acquisition cost.

Evaluate Each Core Switch Separately: While this set of Core switch category tested very well, there are differences between vendors especially in the RFC 2889 congestion, RFC 3918 IP Multicast, power consumption and cloud simulation test. Therefore, it is recommended to review each supplier's results and make purchase decisions accordingly.

Deploy Core Switches that Demonstrate Efficient Power and Cooling: In addition to solid TEER values and low WATTSATIS , all Core switches, except the EX8216, tested support front-to-back or rear-to-front cooling in support of data center hot/cold aisle airflow designs. Juniper does support front to back airflow via the purchase of a 3rd party adapter. Therefore, it is recommended that these Core switches can be used as part of an overall green data center initiative.

Ready for Storage Enablement: Most of the Core switches demonstrated the performance and latency required to support storage enablement or converged I/O. In fact, all suppliers have invested in storage enablement such as support for CEE, FCoE, iSCSI, NAS, etc., and while these features were not tested in the Lippis/Ixia evaluation, these switches demonstrated that the raw capacity is built into the switches for its support.

Evaluate Core Switches for Network Services and Fabric Differences: There are differences between suppliers in terms of the network services they offer as well as how their Core switches connect to ToR/Leaf switches to create a data center fabric. It is recommended that IT business leaders evaluate Core switches with ToR switches to assure that the network services and fabric attributes sought after are realized.

Connect Servers at 10GbE: The Lippis/Ixia test demonstrated the performance and power consumption advantages of 10GbE networking, which can be put to work and exploited for corporate advantage. For new server deployments in private/public data center cloud networks, 10GbE is recommended as the primary network connectivity service as a network fabric exists to take full advantage of server I/O at 10GbE bandwidth and latency levels.

40GbE Uplink Termination Options Become Available: The OmniSwitch 10K, Juniper EX8216 and Arista 7504 possess the hardware architecture and scale to support 40GbE and 100GbE modules. Some of these vendors are shipping such modules and should be considered especially to terminate ToR switches with 40GbE uplinks. While the Lippis/Ixia test did not evaluate 40GbE or 100GbE, we do demonstrate that Core switch hardware platforms scale to support the highest 10GbE port density possible with the vendor's current highest 10GbE port density modules. Therefore, trading off four 10GbE ports for a 40GbE port would not concern us in terms of Core switch performance.

© Lippis Enterprises, Inc. 2011

Consider a Two-Tier Network Architecture: With Core switch 10GbE port densities in the 128 to 256 port range forwarding L2 and L3 packets at wire speed at nano and μs latency, a two-tier network architecture made up of Core and ToR is preferable for private/public data center cloud computing scale deployments. The OmniSwitch 10K, Juniper EX8216 and Arista 7504 possess the performance and power efficiency required to deploy a two-tier network architecture.

Evaluate Link Aggregation Technologies: Not tested in the Lippis/Ixia evaluation was a Core switch's ability to support MC-LAG (Multi-Chassis Link Aggregation Group), TRILL (Transparent Interconnection of Lots of Links) or SPB (802.1aq (Shortest Path Bridging). This is a critical test to determine Ethernet fabric scale. It is recommended that IT business leaders and their network architects evaluate each supplier's MC-LAG, TRILL and SPB approach, performance, congestion and latency characteristics.

© Lippis Enterprises, Inc. 2011

Terms of Use

This document is provided to help you understand whether a given product, technology or service merits additional investigation for your particular needs. Any decision to purchase a product must be based on your own assessment of suitability based on your needs. The document should never be used as a substitute for advice from a qualified IT or business professional. This evaluation was focused on illustrating specific features and/or performance of the product(s) and was conducted under controlled, laboratory conditions. 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.

Reasonable efforts were made to ensure the accuracy of the data contained herein but errors and/or oversights can occur. 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 vendors 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. Accordingly, this document is provided "as is," and Lippis Enterprises, Inc. (Lippis), gives no warranty, representation or undertaking, whether express or implied, and accepts no legal responsibility, whether direct or indirect, for the accuracy, completeness, usefulness or suitability of any information contained herein.

By reviewing this document, you agree that your use of any information contained herein is at your own risk, and you accept all risks and responsibility for losses, damages, costs and other consequences resulting directly or indirectly from any information or material available on it. Lippis Enterprises, Inc. is not responsible for, and you agree to hold Lippis Enterprises, Inc. and its related affiliates harmless from any loss, harm, injury or damage resulting from or arising out of your use of or reliance on any of the information provided herein.

Lippis Enterprises, Inc. makes no claim as to whether any product or company described herein is suitable for investment. You should obtain your own independent professional advice, whether legal, accounting or otherwise, before proceeding with any investment or project related to any information, products or companies described herein. When foreign translations exist, the English document is considered authoritative. To assure accuracy, only use documents downloaded directly from www.lippisreport.com .

No part of any document may be reproduced, in whole or in part, without the specific written permission of Lippis Enterprises, Inc. All trademarks used in the document are owned by their respective owners. You agree not to use any trademark in or as the whole or part of your own trademarks in connection with any activities, products or services which are not ours, or in a manner which may be confusing, misleading or deceptive or in a manner that disparages us or our information, projects or developments.

About Nick Lippis

Nicholas J. Lippis III is a world-renowned authority on advanced IP networks, communications and their benefits to business objectives. He is the publisher of the Lippis Report, a resource for network and IT business decision makers to which over 35,000 executive IT business leaders subscribe. Its Lippis Report podcasts have been downloaded over 160,000 times; i-Tunes reports that listeners also download the Wall Street Journal's Money Matters, Business Week's Climbing the Ladder, The Economist and The Harvard Business Review's IdeaCast. Mr. Lippis is currently working with clients to design their private and public virtualized data center cloud computing network architectures to reap maximum business value and outcome.

He has advised numerous Global 2000 firms on network architecture, design, implementation, vendor selection and budgeting, with clients including Barclays Bank, Eastman Kodak Company, Federal Deposit Insurance Corporation (FDIC), Hughes Aerospace, Liberty Mutual, Schering-Plough, Camp Dresser McKee, the state of Alaska, Microsoft, Kaiser Permanente, Sprint, Worldcom, Cigitel, Cisco Systems, Hewlett Packet, IBM, Avaya and many others. He works exclusively with CIOs and their direct reports. Mr. Lippis possesses a unique perspective of market forces and trends occurring within the computer networking industry derived from his experience with both supply and demand side clients.

Mr. Lippis received the prestigious Boston University College of Engineering Alumni award for advancing the profession. He has been named one of the top 40 most powerful and influential people in the networking industry by Network World. TechTarget an industry on-line publication has named him a network design guru while Network Computing Magazine has called him a star IT guru.

Mr. Lippis founded Strategic Networks Consulting, Inc., a well-respected and influential computer networking industry-consulting concern, which was purchased by Softbank/Ziff-Davis in 1996. He is a frequent keynote speaker at industry events and is widely quoted in the business and industry press. He serves on the Dean of Boston University's College of Engineering Board of Advisors as well as many start-up venture firm's advisory boards. He delivered the commencement speech to Boston University College of Engineering graduates in 2007. Mr. Lippis received his Bachelor of Science in Electrical Engineering and his Master of Science in Systems Engineering from Boston University. His Masters' thesis work included selected technical courses and advisors from Massachusetts Institute of Technology on optical communications and computing.

© Lippis Enterprises, Inc. 2011