



Cisco Nexus® 3548 Top-of-Rack Switch Performance and Power Test



A Report on the:

Cisco Nexus® 3548 Top-of-Rack Switch

July, 2013

Acknowledgements

We thank the following people for their help and support in making possible this first evaluation of the fastest 10GbE data center top-of-rack Ethernet switch:

Victor Alston, CEO of Ixia—for his continuing support of these industry initiatives throughout its execution since 2009.

Jacob Rapp, Manager, Technical Marketing at Cisco and Nicolas Delecroix, Technical Marketing at Cisco for their help and engineering support.

Scott O. Bradner of Harvard University for his work at the IETF 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.

Siemon for the use of copper and fiber optic cables equipped with QSFP+ connectors to link Ixia test equipment to 40GbE 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.

Jim Smith, VP of Marketing at Ixia, for his support and contributions to creating a successful industry event.

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.

License Rights

© 2013 Lippis Enterprises, Inc. All rights reserved. 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 for informational purposes and may not 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 viewer 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.

Cisco Systems Nexus® 3548

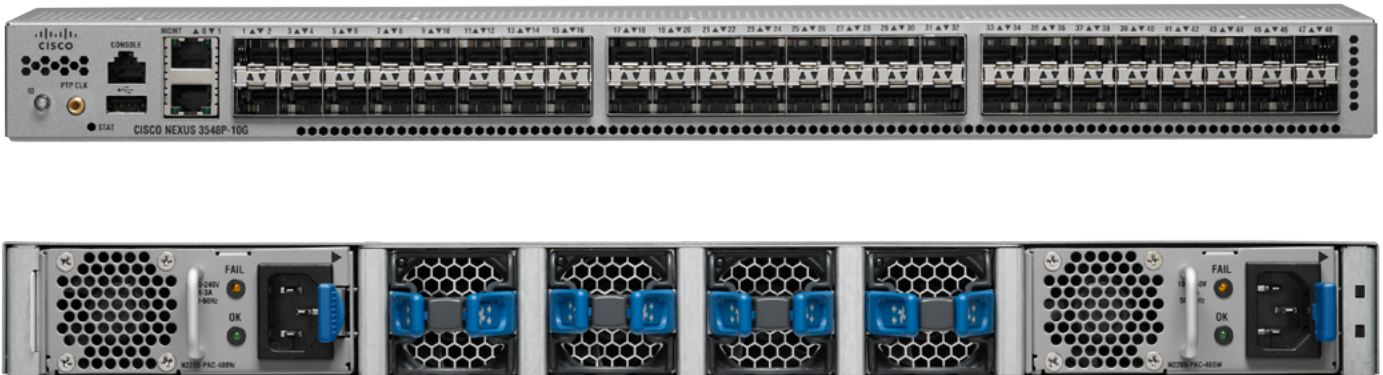
Cisco Systems launched its new entry into the enterprise data center and cloud computing markets with its new Nexus® 3548 switch in September 2012. This new Top-of-Rack switch pushes the envelope on performance, power consumption, density, and introduces new parallelized switching logic.

The Nexus® 3548 switch is a purpose-built Top-of-Rack (ToR) switch designed to support emerging 1/10/40 GbE enabled servers in enterprise/cloud data centers and high frequency trading scenarios. The Nexus® 3548 provides 1, 10 and 40GbE physical server connections to ease transition to high-performance server connections while providing unique features for virtualized data center environments. The NX-OS operating system, with its modularity and highly scalable design, provides a platform where zero-impact operations become a reality.

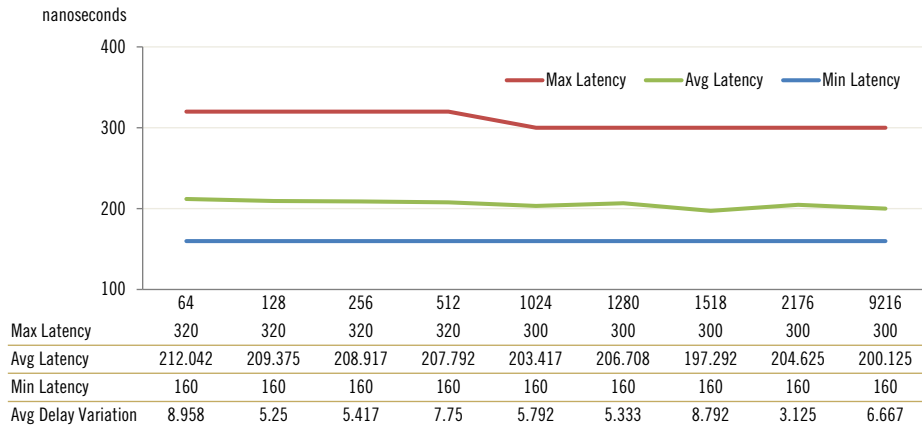
Cisco Nexus® 3548 Test Configuration			
	Hardware	Software Version	Port Density
Device under test	Cisco Nexus® 3548 http://www.cisco.com/en/US/products/ps12581/index.html	5.0(3)A1(2)	48
Test Equipment	Ixia XG12 High Performance Chassis Xcellon Flex AP10G16S (16 port 10G module) Xcellon Flex Combo 10/40GE AP (16 port 10G / 4 port 40G) http://www.ixiacom.com/	IxOS 6.40 EA IxNetwork 7.0 EA &	
Cabling	Optical SFP+ connectors. Laser optimized duplex lc-lc 50 micron mm fiber, 850nm SFP+ transceivers www.leviton.com OM3 multimode 2 meter		

Cisco Systems submitted its Nexus® 3548 ToR switch into the Lippis/Ixia test. The Nexus® 3548 is a member of Cisco's Nexus® family of ToR switches capable of supporting 48 1/10GbE ports in a small 1RU footprint. The Nexus® 3548, in combination with its breakthrough Algo Boost and Warp Mode technology, offers unprecedented latency reduction by parallelizing switching logic, storing different data in the same tables and

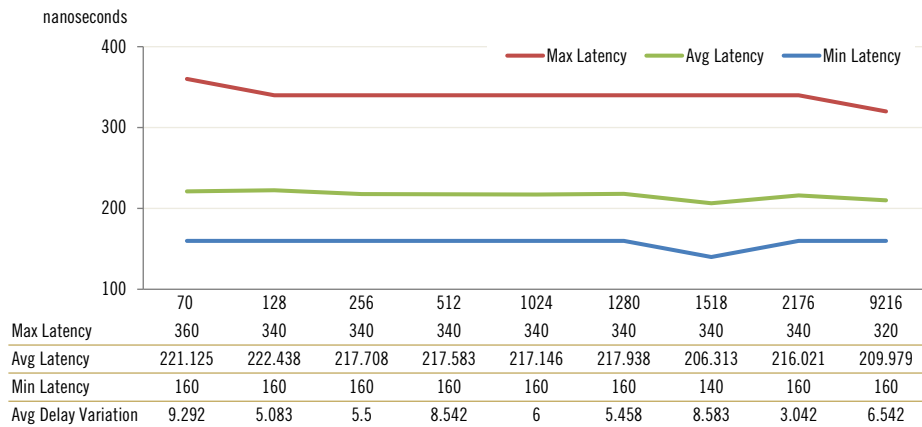
lastly, by building network congestion monitoring into the switching ASIC or application-specific integrated circuit. To understand the capabilities of the Cisco Nexus® 3548, we tested its performance in throughput, packet loss, latency for unicast, multicast plus stateful and stateless traffic flows. To understand its energy conservation claims, we tested for power consumption.



Cisco Nexus® 3548 RFC2544 Layer 2 Latency Test



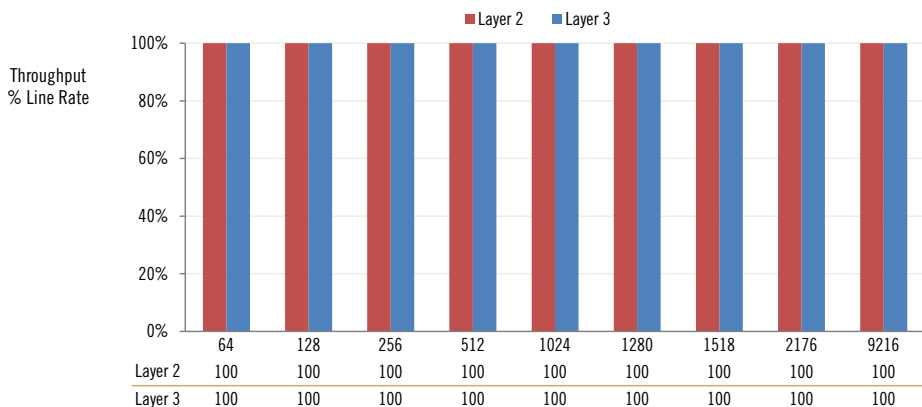
Cisco Nexus® 3548 RFC2544 Layer 3 Latency Test



We populated and tested the Cisco Systems Nexus® 3548 with 48 10GbE ports, which is its full capacity. Its average latency was the best we have ever measured and ranged from a low of 197 ns to a high of 212 ns for layer 2 traffic. Its average delay variation ranged between 5.25 and 8.958 ns, providing consistent latency across all packet sizes at full line rate.

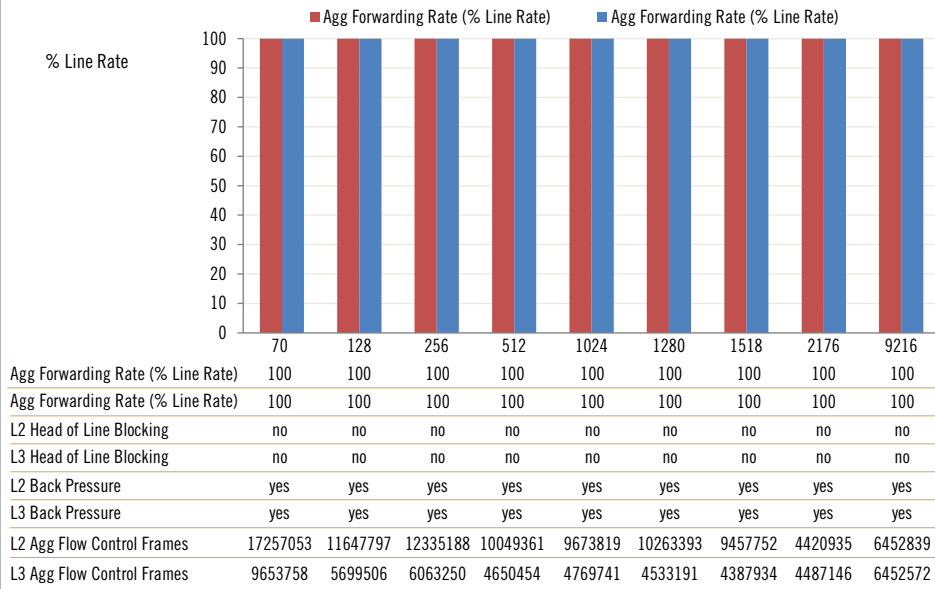
For layer 3 traffic, the Cisco Systems Nexus® 3548 measured average latency ranged from a low of 206 ns at 1518 Bytes to a high of 222 ns at 128 Byte size packet, which is the fastest ToR switch we have ever measured. Its average delay variation for layer 3 traffic ranged between 3 and 9 ns, providing consistent latency across all packet sizes at full line rate.

Cisco Nexus® 3548 RFC 2544 L2 & L3 Throughput Test

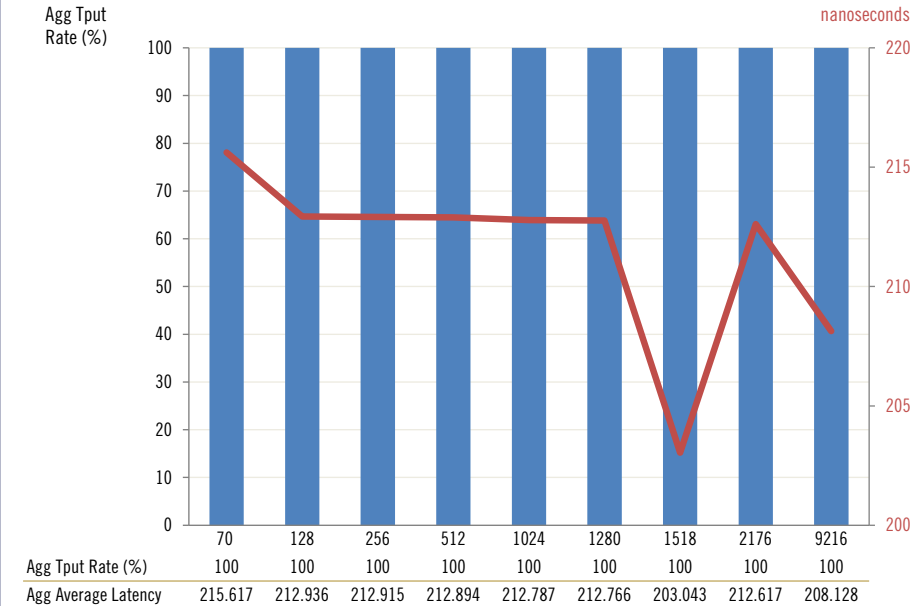


The Cisco Nexus® 3548 demonstrated 100% throughput as a percentage of line rate across all 48 10GbE ports. In other words, not a single packet was dropped while the Cisco Nexus® 3548 was presented with enough traffic to populate its 48 10GbE ports at line rate simultaneously for both L2 and L3 traffic flows.

Cisco Nexus® 3548 RFC 2889 48-port Congestion Test



Cisco Nexus® 3548 RFC 3918 Multicast Test



A 48-port 10GbE congestion test stressed the Nexus® 3548's switch processing and buffer architecture's congestion management subsystems. Twelve groups each consisting of four 10GbE ports per were configured for this congestion test. In each group, one receiver was presented with two times its stated capacity of 10GbE. The Nexus® 3548 demonstrated 100% of aggregated forwarding rate as a percentage of line rate during congestion conditions across all 10GbE ports. There was no Head of Line Blocking or HOLB observed which means that as a 10GbE port on the Nexus® 3548 became congested, it did not impact the performance of other switch ports. During the congestion test, the Nexus® 3548 ran NXOS version was 6.0(2)A1(1) operating system, which does support Link Level Flow Control on the Nexus® 3548. Therefore, back pressure or control/pause frames were detected by Ixia test gear signaling it to slow down the rate of incoming traffic flow to mitigate congestion, which is an industry best practice.

The Cisco Nexus® 3548 demonstrated 100% aggregated throughput for IP multicast traffic with latencies ranging from a low of 203 ns at 1518 Byte size packet to a high of 215 ns at 68 Byte size packets. These are the fastest IP multicast forwarding measurements we have observed at the Ixia iSimCity lab during the Lippis/Ixia tests.

Note that we report zero packet loss at wire speed for unicast and multicast tests. Actual measurements were taken both at 100% and 99.9% of line rate as when the Nexus® 3548 runs at 100% line rate with ppm set to 100, latencies are slightly higher by some 20 ns. However,



Cisco Nexus® 3548 Power Consumption Test	
Watts _{ATIS} /10GbE port	4.81
3-Year Cost/Watts _{ATIS} /10GbE	\$17.59
Total power cost/3-Year	\$844.10
TEER Value	197
Cooling	Front to Back, Reversible

when we change the ppm to 0 at 100% of line rate, the Ixia clock is overrunning the Nexus® 3548 clock, causing measurement issues. Therefore, setting the ppm to 0 at 99.9% line rate mitigated the clocking issue and produced the results reported here. There's currently an Ixia anomalous behavior where setting 100ppm adjustment with any rate, not only 100%, results in a reported latency increase of 20ns average. Ixia support engineers agree that this is a problem and are in review.

The Cisco Nexus® 3548 represents a new breed of data center ToR switches with power efficiency being a core value. Its Watts_{ATIS}/port is \$4.81, and TEER value is 197. Note that the total number of Nexus® 3548 10GbE ports to calculate Watts/10GbE were 48.

The Cisco Nexus® 3548 power cost per 10GbE is calculated at \$5.86 per year. The three-year cost to power the Nexus® 3548 is estimated at \$844.10 and represents approximately 2% of its list price, which is better than industry average by some 20%. Industry three-year power cost as a percentage of list pricing is 2.47%. Keeping with data center best practices, its cooling fans flow air front-to-back and/or back-to-front.



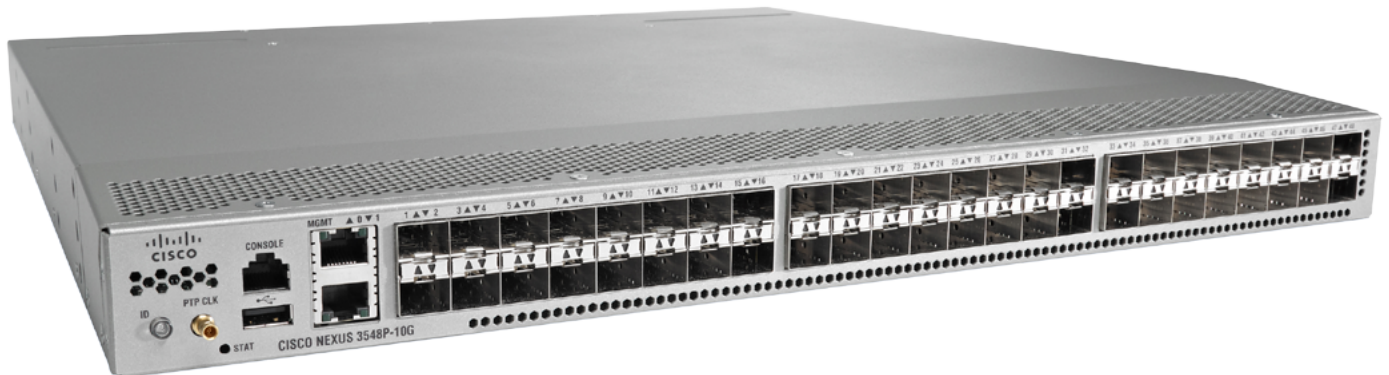
Discussion

The Cisco Systems Nexus® 3548 is the fastest ToR switch that we have tested at these Lippis/Ixia tests by a large amount. The Nexus® 3548 forwards packets in slightly more than half the time of the next fastest switch we have tested! That is, it's nearly twice as fast as the fastest previous switch tested. For IP multicast, we find that packets take some 60% longer to flow through the fastest previous switch tested than the Nexus® 3548. That is, the Nexus® 3548 is the fastest IP multicast forwarding switch we have tested to date, being able to forward packets some 62.5% faster than the previously fastest ToR switch tested. The Nexus® 3548 is an engineering achievement for both its raw packet processing performance, congestion management and value added Algo Boost and Warp Mode technology.

What is striking about the Nexus® 3548 is that its built upon custom ASICs at a time in the industry when merchant silicon is touted as allegedly superior. The Nexus® 3548 proves that hypothesis unfounded.

The Nexus® 3548 is competitive with other ToR switches from a price, performance, latency and power consumption point of view. The Nexus® 3548, when combined with its Algo Boost and Warp Mode technologies, offers a powerful ToR solution for high frequency trading, high performance computing, cloud computing infrastructure and high performance data center network requirements.

The Cisco Systems Nexus® 3548 was found to have best-in-class network performance, including the lowest L2/L3 latency for both unicast and multicast traffic ever observed in a Lippis/Ixia test. In addition to these breakthroughs, the Nexus® 3548 runs on low power, offers flexible cooling options and packaged in a 1RU compact form factor. Cisco Systems Nexus® 3548 is designed to meet the requirements for private/public cloud networks and especially, high frequency trading. Based upon these Lippis/Ixia tests, it achieves its design goals handsomely.



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 L2 unicast, L3 unicast and L3 multicast traffic was measured for packet sizes of 64, 128, 256, 512, 1024, 1280, 1518, 2176 and 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 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.

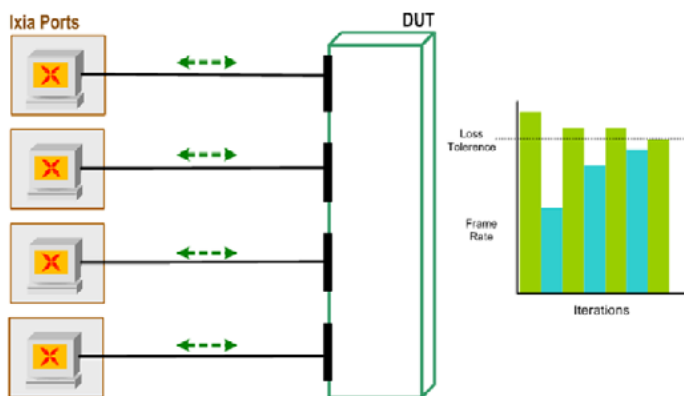
To determine latency, frames are transmitted for a fixed duration. Frames are tagged once in each second and during half of the transmission duration, then tagged frames are transmitted. The receiving and transmitting timestamp on the tagged frames are compared. The difference between

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 2544 Throughput/Latency



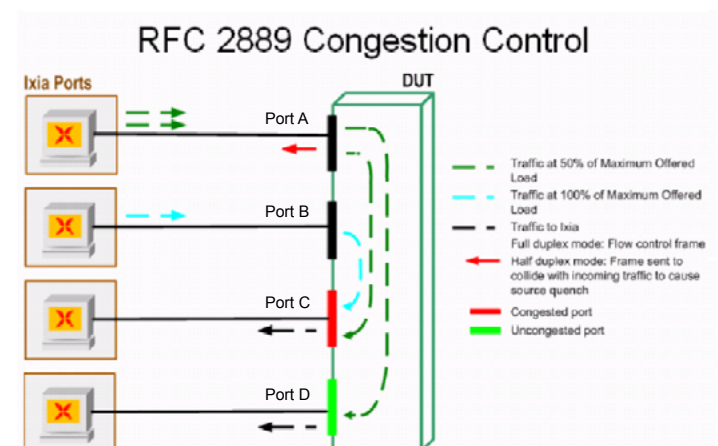
RFC 2889 Congestion Control Test

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. 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 Internet Group Management Protocol (IGMP), Protocol Independent Multicast (PIM), receiver parameters (group addressing), source parameters (emulated PIM routers), frame sizes, initial line rate and search type.

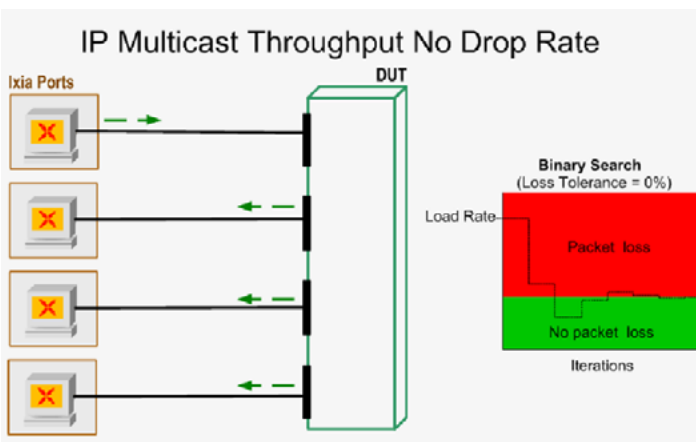
Test Methodology: This test calculates the maximum DUT throughput for IP Multicast traffic using either a binary or a linear search, and to collect Latency and Data

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 (Open Shortest Path First) or ISIS (Intermediate System-Intermediate System) as IGP (Internet Gateway Protocol) 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) 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:

$$WATIS = 0.1 * (\text{Power draw at 0\% load}) + 0.8 * (\text{Power draw at 30\% load}) + 0.1 * (\text{Power draw at 100\% load}).$$

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 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.

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

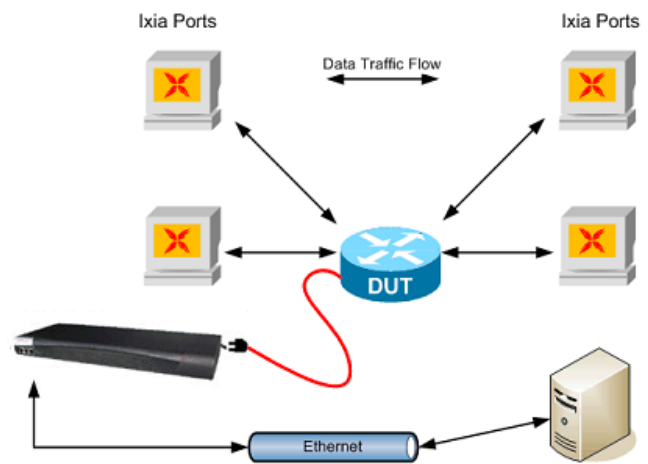
$3 * 365 * 24 = 3 \text{ years @ } 365 \text{ days/yr @ } 24 \text{ hrs/day}$

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.



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 is not responsible for, and you agree to hold Lippis 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 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. 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 200,000 times; iTunes 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. He is also the co-founder and conference chair of the Open Networking User Group, which sponsors a bi-annual meeting of over 200 IT business leaders of large enterprises. Mr. Lippis is currently working with clients to design their private and public virtualized data center cloud computing network architectures with open networking technologies 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, 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 firms' 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.