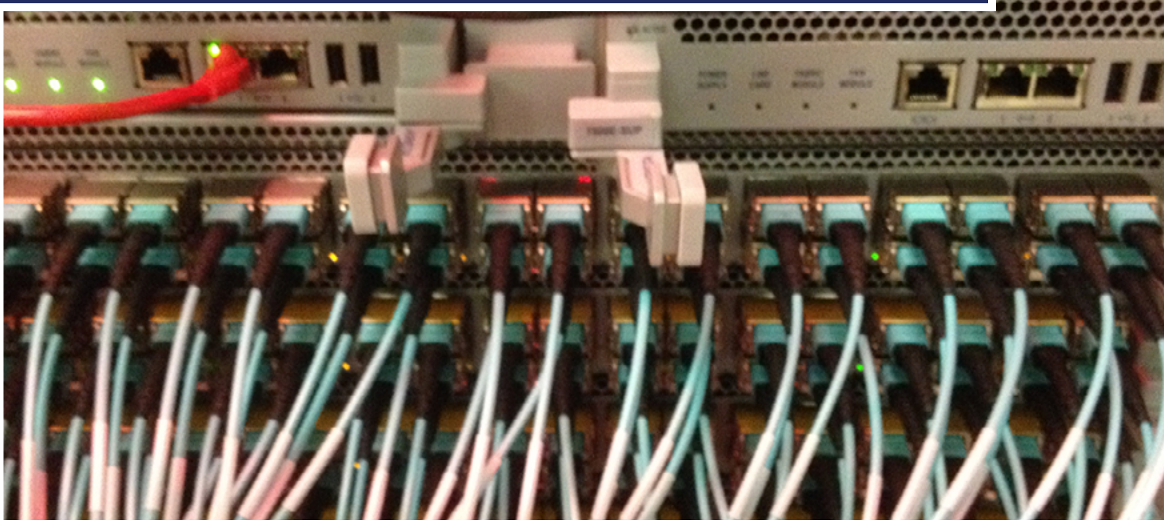
Arista 7500E Software-Defined Cloud Network Switch Performance Test



A Report on the:

Arista 7500E
Software-Defined Cloud Network Switch

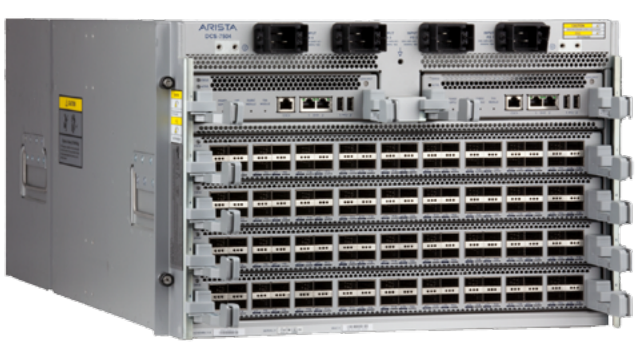
January 2014

The Arista 7500E Modular Data Center Switch

In this Lippis test report, we review the Arista 7500E, and drawing from our experience testing/evaluating numerous switches from companies throughout the industry, Arista's 7500E is positioned to solidly address the scale and density issues that plague the modern data center. In terms of performance, the 7500E boasts line rate throughput at 10/40/100GbE densities and deep packet buffers not available from other suppliers. Arista's motto seems to be "in standards we trust" as its key two-tier architecture protocols are L2 Multi-Chassis Link Aggregation (MLAG) and L3 Equal Cost MultiPath (ECMP) with Virtual eXtensible LAN (VXLAN) for network virtualization support.

The 7500E is the second generation of the 7500 Series, and provides backwards compatibility with the first generation linecards, management modules and fabrics providing investment protection for a system that scales from 10Tbps in the first generation to over 30Tbps in the 7500E Series. According to Arista, Investment protection ensures customers can upgrade in place and take advantage of the expanded series of linecards, performance and deeper buffers.

Its engineers designed the 7500E with the deepest buffers in data center switches to manage congestion and deliver consistently low latency under load. The 7500E has a revolutionary design, utilizing integrated optics that enables the industry's highest density triple speed 10/40/100GbE system that dramatically cuts down cost per port. At 4 watts per 10GbE port, the 7500E consumes less power than a Christmas tree light bulb on a 10GbE/port basis. Economically, the 7500E boasts 10GbE ports being half of current prices with its 40/100GbE ports setting a new low-cost standard. With these huge improvements in deterministic performance, system density and economics, Arista is enabling the era of cloud networking data centers with the power to create massively scalable networks, ranging from 100s to 100,000s of wire-speed 10GbE+ interfaces, that can run modern data center applications, such as big data/Hadoop, mass virtualization, SDN, large-scale Web and high performance compute (HPC).



The 7500E design offers a superior internal fabric architecture that is cell-based, rather than hashing-based. As a result, the switch can handle all workloads in a modern data center, irrespective of traffic flows and even under failure such as a failed fabric module. This enables the 7500E to build an application-agnostic Universal Cloud Network. Packets are stored on the ingress ports in Virtual Output Queues (VOQ) with associated deep packet buffers. Once the egress port is ready to send data, packets are sent from ingress port to egress port through the internal fabric—in essence, a credit-based, scheduled fabric architecture. Typically, these high-end attributes result in compromised port density, but its the combination of the 7500E's architectural benefits and industry-leading port density that makes the Arista 7500E stand out.

Key highlights of the Arista 7500E are:

- One of highest density spine switches in the industry – 1,152x10G, 288x40G, 96x100G in 11RU
- Scalable architecture for all applications: cell-based fabric, Virtual Output Queues (VOQ), deep packet buffers (125 msec per port which scales across port speeds 10G-40G-100G) (125MB/10G, 500MB/40G, 2GB/100G) by default and 144 GB per chassis.
- 64-way ECMP for large-scale Layer 3 designs
- Active/Active, standards-based MLAG with up to 64 ports per LAG
- VXLAN support for multi-tenant virtualized designs
- Front-to-rear airflow, and power efficient 10/40/100GbE
- Integrated, cost effective, 100GbE optics
- High quality, with measured MTBF of over 250,000 hours
- Redundant and hot-swappable fabric modules, power supplies and supervisors plus N+1 redundant fans. Failure of a fabric does not result in loss of forwarding capacity, unlike switches of similar density.

According to Arista, the 7500E sports switching capacity of 30 Tb per second to support port densities of 1,152 10GbE ports, 288 40GbE ports or 96 100GbE ports. The Arista 7500E has an impressive 256K entry table for IPv4 / IPv6 Host Routes, MAC addresses and ARP tables, which are huge and can support the highest of highly virtualized environments. As with the previous 7500 Series modular switch generation, the 7500E is a family of modular switches available in both four-slot (7 RU) and eight-slot (11 RU) form factors, which are now able to support an unprecedented range of line card options. The 7500E boasts a range of 4 different line card options to provide interface flexibility plus mix and match agility.

The 7500E runs on Arista's Extensible Operating System (EOS), a foundational element of Arista's Software-Defined Cloud Networking (SDCN), unifying the network infrastructure and enabling it to interact with other applications in the data center. Ubiquitous workload mobility, a powerful programmable operating system and visibility are the hallmarks of Arista's unified data center design. Leveraging open standards, Arista can connect 100,000+ hosts/servers in a resilient and scalable topology, leveraging MLAG and ECMP.

Functionality	EOS Features
Automation	ZTP, Chef, Puppet, Ansible, Smart Systems Upgrade
Extensibility and Programmability	Linux, eAPIs (JSON), Python, XMPP, Advanced Event Manager, SqlLite Databases
Application and Network Visibility	MapReduce Tracer, VMTracer, PathTracer, Health Tracer, Latency Analyzer
Network Telemetry	DANZ, sFlow, Splunk Forwarder, Time Stamping
Network Virtualization	VxLAN, L2/L3 Gateway Services integration, VM Tracer for VxLAN
Controller Integration	NSX, vCenter, Openflow, Openstack

Arista Networks EOS provides a differentiated feature set that allows for dynamic zero touch provisioning/replacement (ZTP/ZTR) and real-time streaming network telemetry (LANZ+/sFlow and forwarding to partners like Splunk/ExtraHop). In addition, an onboard database tracks all changes to the forwarding state of the data plane, plus powerful aggregation tools allow easy connectivity of analytics tools. Arista enables network-wide virtualization through VXLAN, offers a fully programmable JSON-based API, and integrates natively with VMware vSphere, Microsoft Object Management Infrastructure (OMI), OpenStack, OpenFlow and other leading provisioning, orchestration and automation platforms to streamline IT operations.

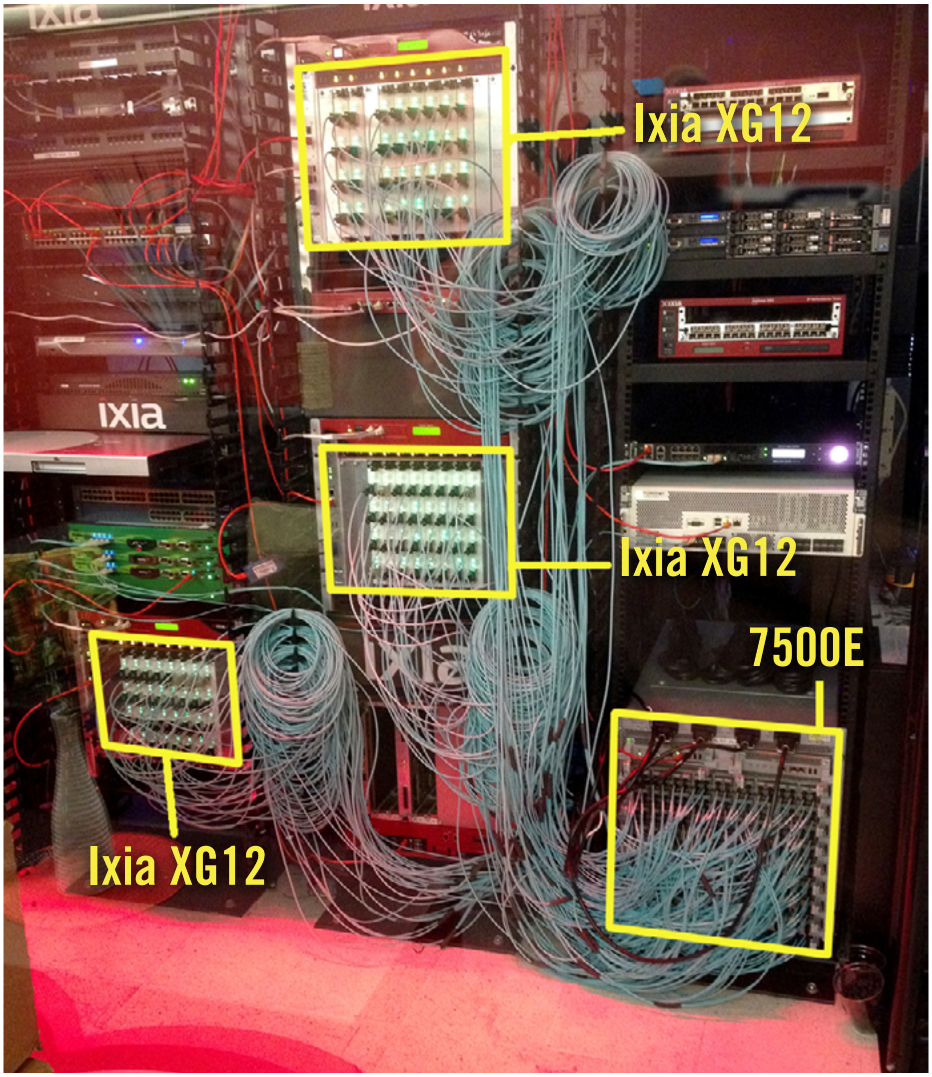
Further, the Arista eAPI leverages industry standard protocols to allow NetOps to send requests to the 7500E and receive structured data in return, which enables integration of EOS with leading orchestration and provisioning tools plus applications. Lastly, the 7500E supports the ability to integrate tightly with backend management systems and other elements of the data center architecture with its smart system upgrade functionality.

While the new features of the Arista 7500E chassis offer countless configuration options, we have run exhaustive tests of the chassis' core competencies in switching latency, throughput, congestion, power consumption and IP multicast to ensure it can perform at the high level that Arista states. What we found is that the Arista 7500E is the most powerful Layer 2 plus Layer 3 data center core switch on the market as of this writing.

Arista 7500E Test Configuration

Only one other data center switch has ever been tested at full 288 40GbE ports, which forced Ixia engineers to design a new approach to test at this scale. To test the Arista 7500E, we connected 288 40GbE ports to three Ixia XG12 High Performance Chassis running its IxOS 6.50 operating system and IxNetwork 7.10 test script suite. To deliver 288 40GbE ports to the Arista 7500E, an Ixia CXP port capable of supplying 120Gbps of network traffic was split into three 40GbE QSFP+ ports. This was accomplished via the Ixia Xcellon-Multis XM40GE12 QSFP+FAN 40GE Load Modules, which support four CXP ports. Eight Ixia Xcellon-Multis load modules populated each XG12 chassis, delivering 96 40GbE streams of line rate traffic. Three XG12s populated with 96 40GbE each delivered 288 40GbE of line rate traffic flow into the Arista 7500E spine switch at varying packet sizes.

The Arista 7500E's 288 40GbE ports were connected via Ixia Multis CXP-to-3-40GE QSFP Active Optical Cable (AOC). The optical wavelength was 850NM, and 3-meter optical cables connected the test gear and the Arista 7500E. The Ixia modules connected to the Arista 7500E via its DCS-7500E-36Q-LC 40GbE line cards that support 36 40GbE QSFP+ ports. Lastly, the Arista 7500E ran its EOS 4.12.6 operating system.



Arista 7500E Modular Data Center Switch

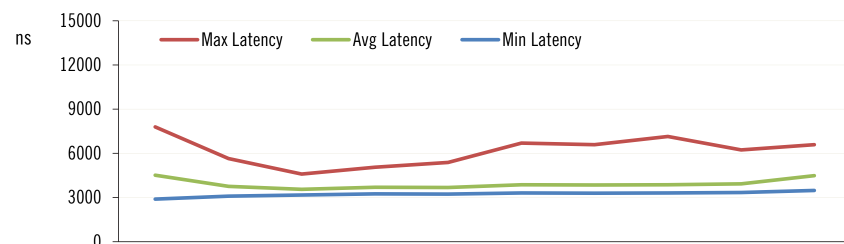
To fully stress the Arista 7500E, we ran RFC 2544 latency and throughput tests on the Arista 7500E, both fully populated at 288 40GbE and 240 40GbE, to understand its line rate or wirespeed throughput capacity levels. These tests were run at both L2 and L3 demonstrating the applicability to large scale cloud environments regardless of topology. The first test was configured in a fully populated 288 40GbE or 1152 10GbE equivalent ports, while the second test was configured for 240 40GbE or 960 10GbE ports. We also tested the Arista 7500 for congestion, IP multicast, power consumption and cloud simulation. We conducted additional tests to understand the effectiveness of its VOQ buffer architecture, hashing algorithm in a 64-way ECMP configuration and 100GbE line card performance. The following details Lippis test results.

The Arista 7500E spine switch was tested across 288 ports of 40GbE in this test at both 50% and 100% of line rate to provide a wide picture of its performance capabilities. At 50% of line rate, the 7500E produced average latency that ranged from a low of 3.5 microseconds (3549.46 ns) to a high of 4.5 microseconds (4477.614 ns) for L2 traffic forwarding across 64 to 9216 byte size packets measured in store-and-forward mode. At 100% of line rate, the 7500E produced average latency that ranged from a low of 3.3 microseconds (3306.29 ns) to a high of 9.1 microseconds (9117.63 ns) for L2 traffic forwarding across 64 and 9216 byte size packets measured in store-and-forward mode. The 7500E demonstrated one

Arista 7500E Test Configurations
288-40GbE or 1152-10GbE, 240-40GbE or 960 10GbE

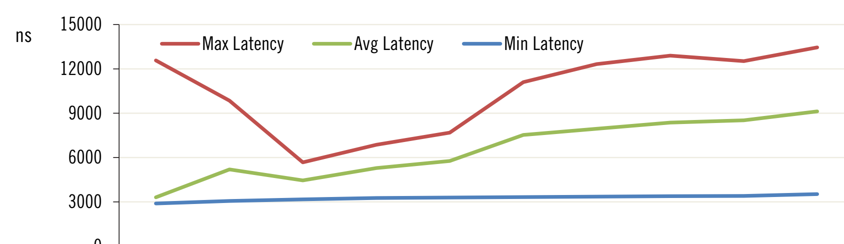
	Hardware	Software Version	Port Density
Device under test	Arista 7500E http://www.aristanetworks.com/en/products/7500series	EOS 4.12.6	288 40GbE or 1152 10GbE 240 40GbE or 960 10GbE
Test Equipment	Ixia XG12 High Performance Chassis	IxOS 6.40 EA IxNetwork 7.10 EASP1	
	(3) Ixia XG12 High Performance Chassis		
	(32) Xcellon-Multis XM40GE12 QSFP+FAN 40GE Load Module		
	http://www.ixiacom.com/		
Cabling	Ixia's Multis CXP-to-3-40GE QSFP Active Optical Cable (AOC)		
	http://www.ixiacom.com/pdfs/datasheets/xcellon-multis-100-40-10-loadmodule.pdf		

Arista 7500E 288 40GbE RFC2544 Layer 2 Latency Test
288-40GbE or 1152-10GbE @ 50% of Line Rate



	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	7785	5635	4587	5052	5385	6692	6582	7147	6222	6592
Avg Latency (ns)	4515.01	3762.50	3549.46	3692.66	3672.06	3871.99	3851.32	3864.37	3928.18	4477.14
Min Latency (ns)	2897	3085	3172	3242	3237	3307	3292	3307	3332	3472
Avg Delay Variation (ns)	33.00	47.77	59.60	76.95	91.62	158.14	158.06	157.37	151.62	392.30

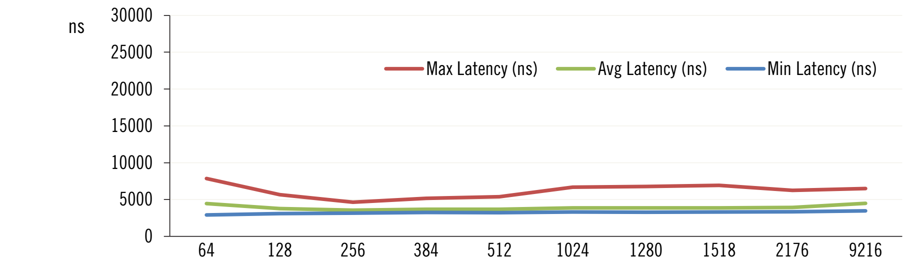
Arista 7500E 288 40GbE RFC2544 Layer 2 Latency Test
288-40GbE or 1152-10GbE @ 100% of Line Rate



	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	12572	9845	5677	6867	7675	11097	12317	12895	12517	13450
Avg Latency (ns)	3306.29	5191.26	4445.00	5286.03	5760.79	7524.54	7952.44	8358.51	8520.19	9117.64
Min Latency (ns)	2895	3067	3172	3267	3287	3330	3360	3385	3407	3520
Avg Delay Variation (ns)	8.00	5.35	11.00	3.00	3.00	3.00	3.44	1.00	3.45	3.00

Arista 7500E 288 40GbE RFC2544 Layer 3 Latency Test

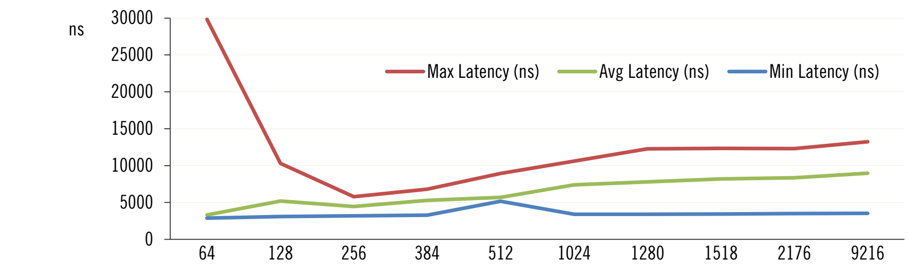
288-40GbE or 1152-10GbE @ 50% of Line Rate



	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	7852	5645	4625	5165	5390	6682	6772	6922	6252	6487
Avg Latency (ns)	4458.86	3757.01	3549.86	3692.85	3671.94	3871.09	3851.45	3863.97	3927.93	4477.21
Min Latency (ns)	2897	3085	3167	3240	3222	3307	3290	3315	3327	3472
Avg Delay Variation (ns)	33.00	47.24	59.30	76.88	91.45	157.41	155.60	157.10	151.78	389.83

Arista 7500E 288 40GbE RFC2544 Layer 3 Latency Test

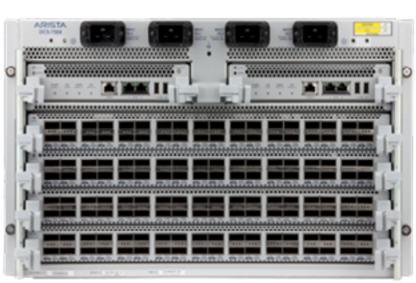
288-40GbE or 1152-10GbE @ 100% of Line Rate

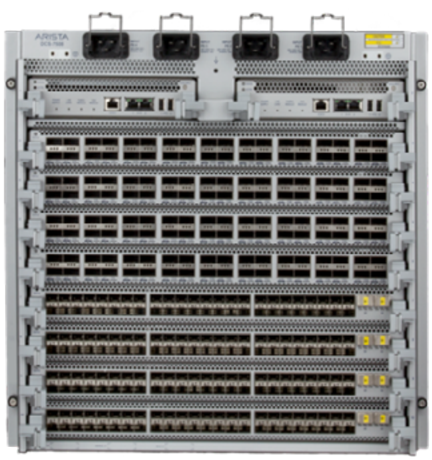


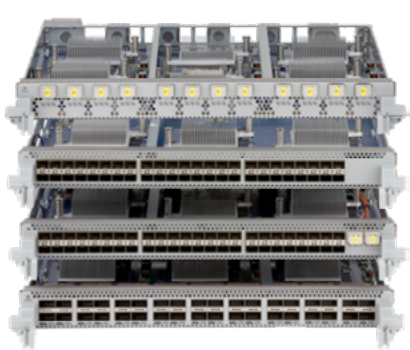
	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	29830	10307	5787	6787	8920	10612	12270	12347	12302	13225
Avg Latency (ns)	3309.26	5200.09	4443.15	5284.00	5688.67	7376.61	7800.31	8198.99	8350.29	8961.48
Min Latency (ns)	2890	3090	3195	3275	5165	3412	3410	3420	3490	3520
Avg Delay Variation (ns)	8.00	4.00	11.00	3.00	3.00	3.00	3.47	1.00	3.44	3.00

of the lowest latency results for a spine switch that we have observed in these Lippis/Ixia test. At 100% line rate, its average delay variation ranged between 1 and 8 ns, providing competitive and consistent latency across all packet sizes at full line rate—a comforting result given the 7500E's design focus on high throughput and low latency critical to cloud networking environments.

During the L3 unicast latency test at 50% line rate populated with 288 40GbE, the Arista 7500E produced average layer 3 store-and-forward latency ranging from a low of 3.5 microseconds (3549.86 ns) to a high of 4.5 microseconds (4477.212 ns) across 64 to 9216 byte size packets. At 100% line rate populated with 288 40GbE, the Arista 7500E produced average L3 unicast store-and-forward latency ranging from a low of 3.3 microseconds (3309.26 ns) to a high of 8.9 microseconds (8961.48 ns) across 64 to 9216 byte size packets. The average delay variation at 100% of line rate ranged between 1 to 11 ns, providing consistent latency across all packet sizes at line rate.

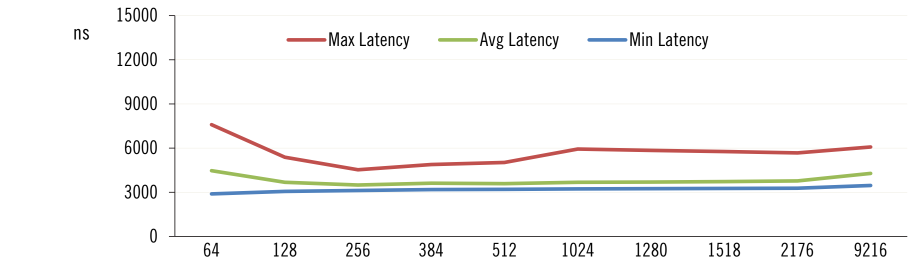






Arista 7500E 240 40GbE RFC2544 Layer 2 Latency Test

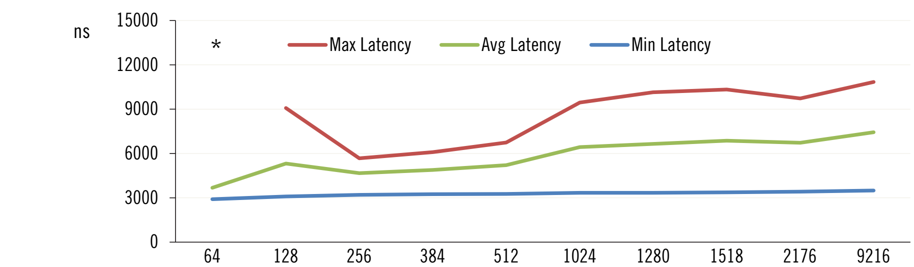
240-40GbE or 960 10GbE @ 50% of Line Rate



	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	7595	5387	4537	4892	5027	5930	5847	5772	5672	6082
Avg Latency (ns)	4460.14	3678.86	3489.40	3610.72	3593.63	3684.56	3691.65	3718.60	3779.44	4287.83
Min Latency (ns)	2895	3067	3127	3185	3195	3225	3247	3267	3272	3460
Avg Delay Variation (ns)	33.00	47.98	54.60	74.44	85.19	133.27	119.01	118.07	111.13	315.17

Arista 7500E 240 40GbE RFC2544 Layer 2 Latency Test

240-40GbE or 960 10GbE @ 100% of Line Rate



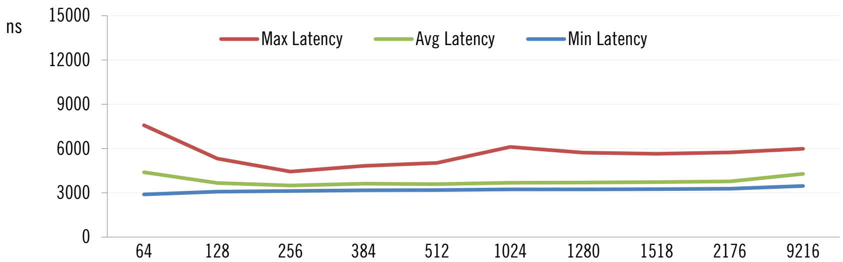
	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	221670*	9072	5675	6090	6742	9452	10145	10327	9725	10835
Avg Latency (ns)	3676.15	5323.39	4663.99	4879.30	5212.79	6428.16	6647.41	6862.00	6717.09	7437.62
Min Latency (ns)	2900	3092	3200	3247	3265	3342	3340	3365	3420	3490
Avg Delay Variation (ns)	3.00	3.00	3.22	3.00	3.00	3.00	3.49	1.00	3.48	3.00

* at the 64Byte packet size maximum latency measurement an anomalistically high recording was observed and not graphed, however its measurement is listed in the table above.

The Arista 7500E spine switch was also tested across 240 ports of 40GbE at 50% and 100% of line rate. At 50% of line rate, the 7500E produced average latency that ranged from a low of 3.5 microseconds (3489.40 ns) to a high of 4.5 microseconds (4460.14 ns) for L2 traffic forwarding across 64 and 9216 byte size packets measured in store-and-forward mode. At 100% of line rate, the 7500E produced average latency that ranged from a low of 3.6 microseconds (3676.15 ns) to a high of 7.4 microseconds (7437.62 ns) for L2 traffic forwarding across 64 and 9216 byte size packets, respectively, measured in store-and-forward mode. At 100% of line rate, the 7500E tested in the 240 40GbE configuration demonstrated lower and more consistent latency than in the 100% line rate 288 40GbE configuration and still one of the lowest latency results for a spine switch that we have observed in these Lippis/Ixia test. Its average delay variation ranged between 1 and 3 ns, providing one of the most consistent latency measurements across all packet sizes at full line rate that we have observed. Low and consistent latency is a fundamental requirement for IP storage, and the Arista 7500E clearly demonstrated its ability to provide InfiniBand-type latency for Ethernet transport of IP storage.

Arista 7500E 240 40GbE RFC2544 Layer 3 Latency Test

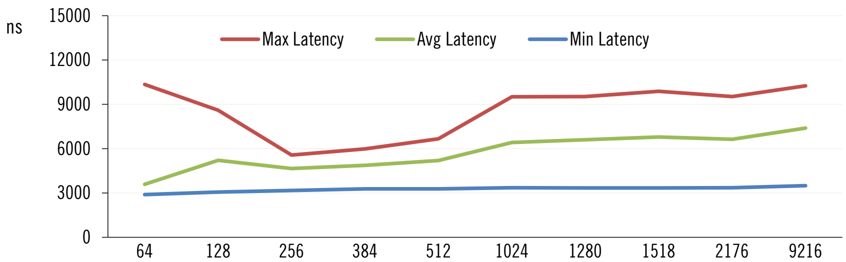
240-40GbE or 960 10GbE @ 50% of Line Rate



	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	7575	5325	4432	4827	5017	6100	5720	5642	5732	5982
Avg Latency (ns)	4392.51	3668.90	3489.16	3610.71	3593.04	3681.36	3690.75	3717.98	3779.14	4287.10
Min Latency (ns)	2897	3080	3125	3175	3192	3227	3237	3250	3280	3457
Avg Delay Variation (ns)	33.00	47.95	54.11	74.30	84.74	131.88	117.70	116.68	110.88	318.52

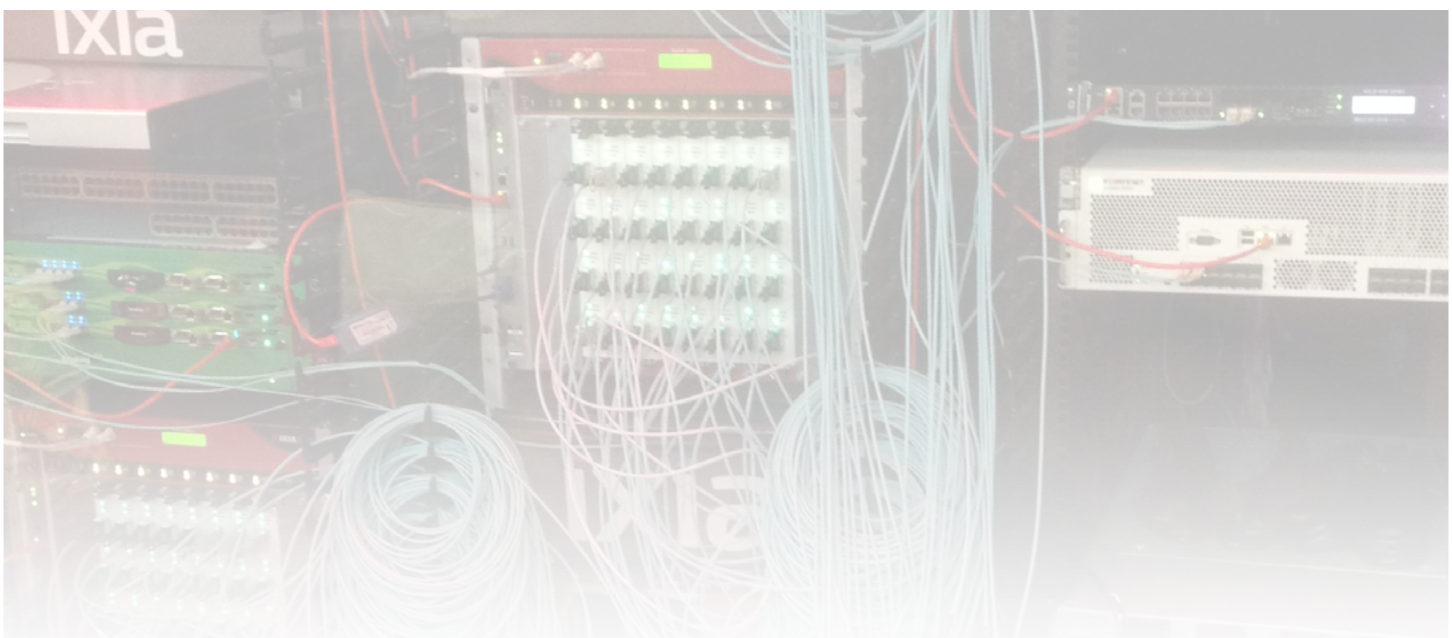
Arista 7500E 240 40GbE RFC2544 Layer 3 Latency Test

240-40GbE or 960 10GbE @ 100% of Line Rate



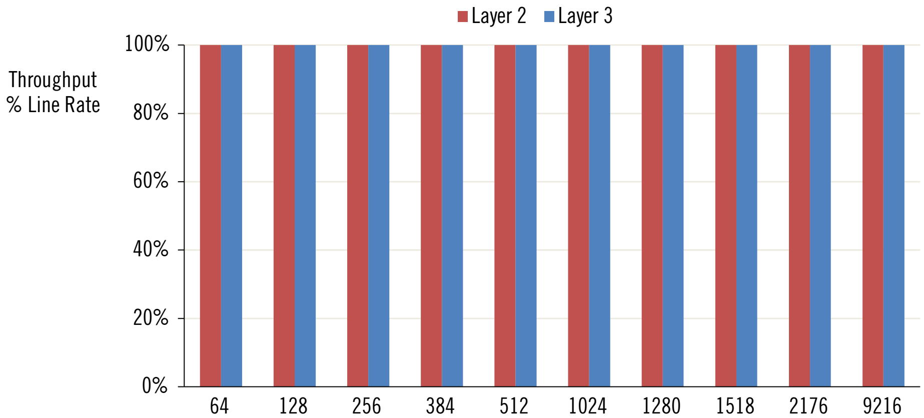
	64	128	256	384	512	1024	1280	1518	2176	9216
Max Latency (ns)	10342	8592	5570	5987	6660	9510	9525	9875	9517	10245
Avg Latency (ns)	3591.24	5201.73	4646.69	4866.11	5197.56	6412.50	6606.19	6785.66	6627.28	7383.57
Min Latency (ns)	2897	3055	3167	3275	3275	3357	3335	3345	3357	3500
Avg Delay Variation (ns)	3.00	3.00	3.23	3.00	3.00	3.00	3.51	1.00	3.54	3.00

For the 240 40GbE L3 unicast traffic store-and-forward latency test at 50% of line rate, the Arista 7500E was observed to produce average latency ranging from a low of 3.5 microseconds (3489.16 ns) to a high of 4.3 microseconds (4392.50 ns) across 64 to 9216 byte size packets. For the 240 40GbE L3 unicast at 100% of line rate traffic store-and-forward latency test, the Arista 7500E was observed to produce average latency ranging from a low of 3.6 microseconds (3591.24 ns) at 64 byte size packets to a high of 7.3 microseconds (7282.57 ns) at 9216 byte size packets. The average delay variation ranged from 1 to 3.5 ns. Just as in the L2 latency test, the 7500E tested in the 240 40GbE configuration demonstrated lower and more consistent latency than the 288 40GbE configuration and still one of the lowest latency results for a spine switch that we have observed in these Lippis/Ixia test. Its average delay variation ranged between 1 and 3.5 ns, providing one of the most consistent latency measurements across all packet sizes at full line rate that we have observed.



Arista 7500E 288-40GbE RFC2544 Throughput Test

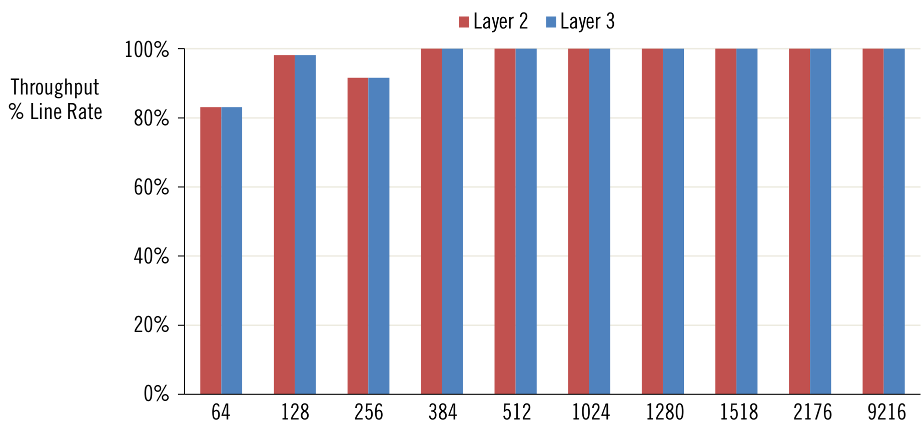
288-40GbE or 1152-10GbE @ 50% of Line Rate



	64	128	256	384	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%

Arista 7500E 288-40GbE RFC2544 Throughput Test

288-40GbE or 1152-10GbE @ 100% of Line Rate

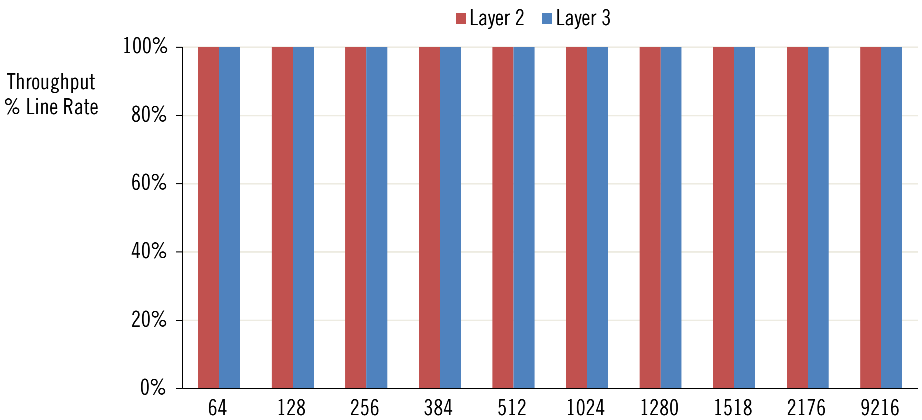


	64	128	256	384	512	1024	1280	1518	2176	9216
Layer 2	83.13%	98.13%	91.56%	100%	100%	100%	100%	100%	100%	100%
Layer 3	83.13%	98.13%	91.56%	100%	100%	100%	100%	100%	100%	100%

At 50% of line rate traffic, the Arista 7500E demonstrated a 100% throughput as a percentage of line rate for all packet sizes. At 100% of line rate traffic, the Arista 7500E demonstrated a range of 83% to 92% throughput as a percentage of line rate for packet sizes of 64 to 256 bytes, and 100% throughput as a percentage of line rate or full wirespeed across all 288 40GbE ports for packet sizes of 384 to 9216 bytes. The 7500E delivers wirespeed performance for all 288 ports of 40GbE with an average packet size of 384 bytes, which closely matches the demands of the highest performance real-world networks.

Arista 7500E 240-40GbE RFC2544 Throughput Test

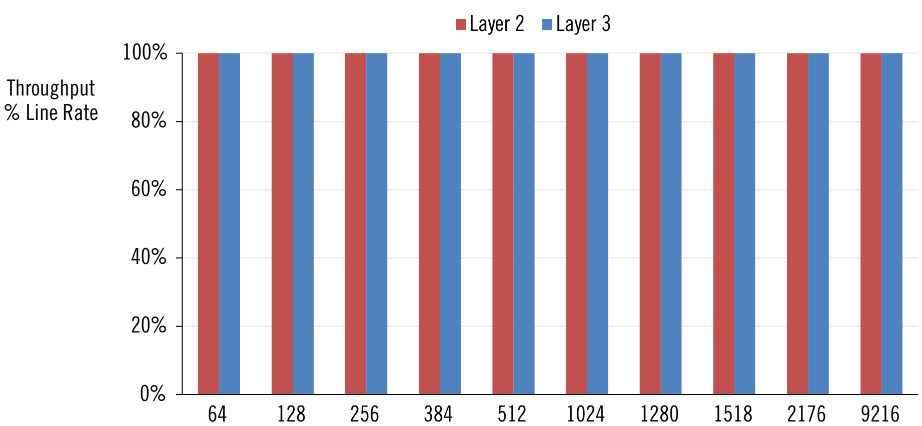
240-40GbE or 960-10GbE @ 50% of Line Rate



	64	128	256	384	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%

Arista 7500E 240-40GbE RFC2544 Throughput Test

240-40GbE or 960-10GbE @ 100% of Line Rate



	64	128	256	384	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%

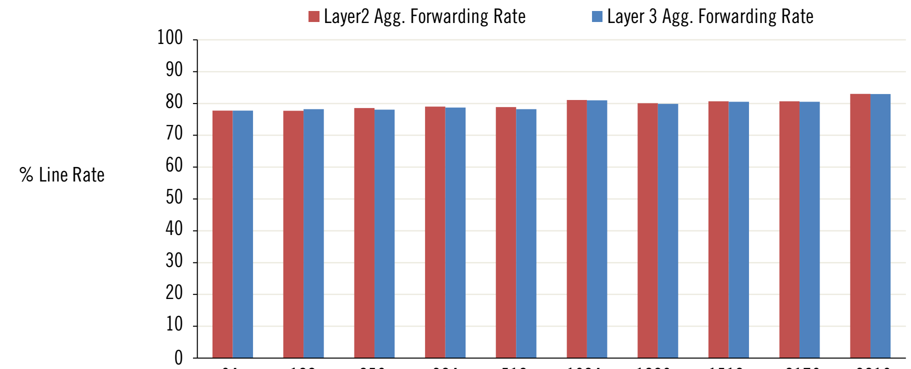
Configured in the 240 40GbE-port density, the Arista 7500E demonstrated 100% throughput as a percentage of line rate or full wirespeed for all packet sizes from 64 to 9216 bytes. In other words, no packets were dropped while the Arista 7500E was presented with enough traffic to populate its highly dense 240 40GbE ports at line rate.

Note that this RFC 2544 throughput test was conducted with 240 40GbE ports or that 30 ports of 40GbE per line card were connected to Ixia test gear. This verifies that the 7500E operates at line rate for all packet sizes when 30 of 36 40GbE ports on each line card are utilized. Each DCS-7500E-36Q-LC 40GbE line card is populated with six Broadcom Dune/ARAD or BCM88650 chips. When the DCS-7500E-36Q-LC line card modules is populated with 30 40GbE ports or five 40GbE ports per ARAD/ BCM88650, we measured full line rate performance across all packet sizes. Otherwise, the DCS-7500E-36Q-LC is line rate above 384 bytes when measured using fixed frame size testing. Note that RFC2544 fixed frame size testing does not necessarily reflect real world traffic profiles hence why the Lippis Cloud Performance iMix test was created.

To understand how the Arista 7500E behaves during periods of congestion, we stressed the 7500E switch processing and buffer architecture's congestion management subsystems. To achieve this goal, one group of four 40GbE ports was configured for congestion testing. A single 40GbE port was flooded at 150% of line rate while Ixia test gear monitored the other three ports. The Arista 7500E demonstrated a range of 77.8% to 83% of aggregated forwarding rate as percentage of line rate during congestion conditions for both L2 and L3 traffic flows. These were the best congestion management numbers we have observed during the Lippis/Ixia test for spine switches, thanks to Arista's adjustable Virtual Output Queuing buffer algorithm plus its generous packet buffers of 18 GB of line card and 144 GB for a fully loaded chassis. There was no Head of Line Blocking (HOLB) observed, which means that as a 40GbE port on the Arista 7500E became congested, it did not impact the performance of other switch ports.

Note that nearly all Arista 7500E congestion results show backpressure detection when, in fact, this is a phantom reading.

Arista 7500E
RFC 2889 Congestion Test



	64	128	256	384	512	1024	1280	1518	2176	9216
Layer 2 Agg. Forwarding Rate	77.76	77.69	78.55	79.05	78.89	81.11	80.10	80.67	80.70	83.02
Layer 3 Agg. Forwarding Rate	77.77	78.24	78.05	78.73	78.19	81.00	79.89	80.56	80.54	82.96
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	no
L2 Agg Flow Control Frames	0	0	0	0	0	0	0	0	0	0
L3 Agg Flow Control Frames	0	0	0	0	0	0	0	0	0	0

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 Arista 7500E large 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 Arista 7500E or Ixia test gear.

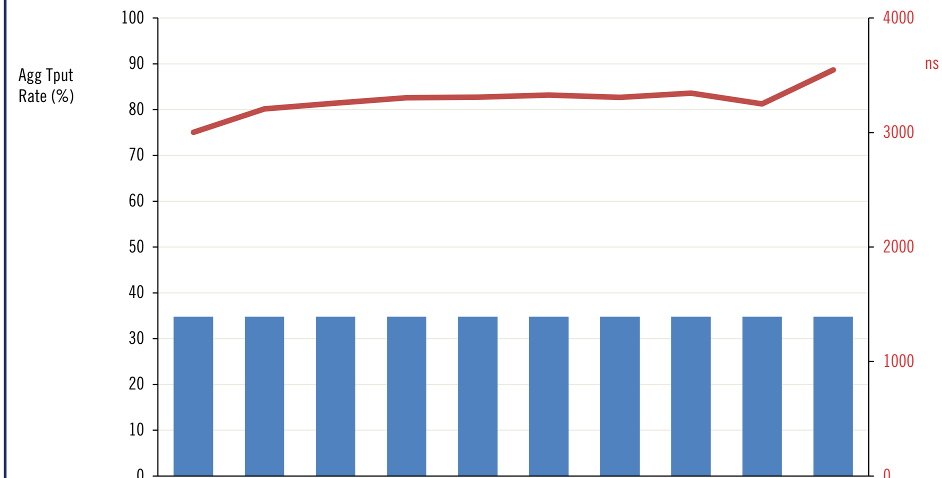
No control/pause frames were detected and observed by Ixia test gear while testing the Arista 7500E, which is the norm in spine switches.

Two types of IP multicast tests were run. They are the accumulated and distributed modes. Accumulated mode runs RFC 3918 as one multicast group where one source transmits to multiple destinations, while in the distributed mode, a single source is distributed evenly to the number of destination ports within a multicast group. The Arista 7500E demonstrated 0.348% aggregated throughput for IP multicast traffic in distributed mode with store-and-forward latencies ranging from a low of 3 microseconds (3002.2 ns) at 64 byte size packets to a high of 3.5 microseconds (3547.13 ns) at 9216 byte size packets. Note that during the "distributed" mode multicast test, 287 groups joined a single, unique multicast group, so that each receive port was only receiving 1/287th, or .348% of the traffic was expected and observed.

In the accumulated IP multicast mode, the Arista 7500E demonstrated nearly 100% of aggregated throughput for IP multicast traffic with store-and-forward latencies, ranging from a low of 3.9 microseconds (3915.81 ns) at 2176 byte size packets to a high of 4.2 microseconds (4184.02 ns) at 9216 byte size packets. As the linecard is wirespeed at 384 bytes and above for all 36x40G ports packet loss of 16.16% and 8.03% was expected and observed at 64 and 256 byte size packets, respectively, during the accumulated IP multicast test.

Arista 7500E RFC 3918 IP Multicast Test

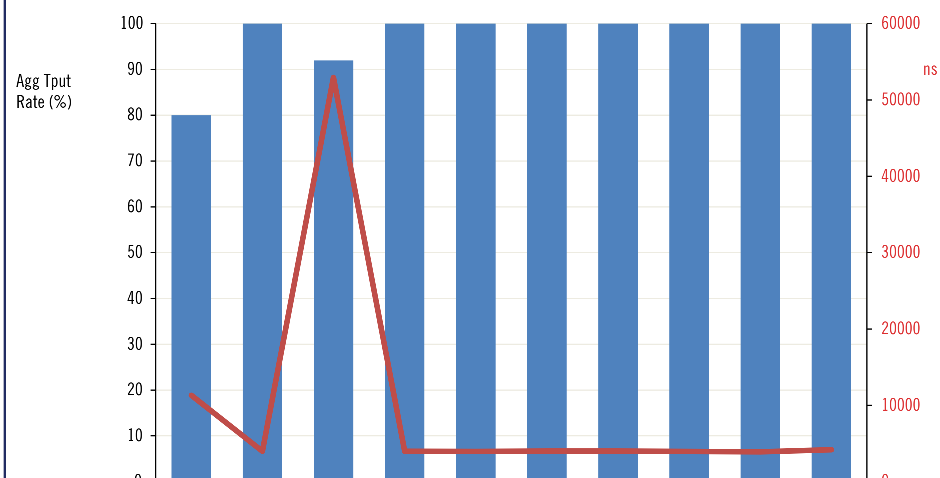
288-40GbE or 1152-10GbE Distributed Multicast Test Mode



	64	128	256	384	512	1024	1280	1518	2176	9216
Agg Tput Rate (%)	34.8	34.8	34.8	34.8	34.8	34.8	34.8	34.8	34.8	34.8
Agg. Average Latency	3002.30	3207.27	3257.06	3304.31	3308.55	3327.08	3307.45	3344.77	3250.00	3547.13

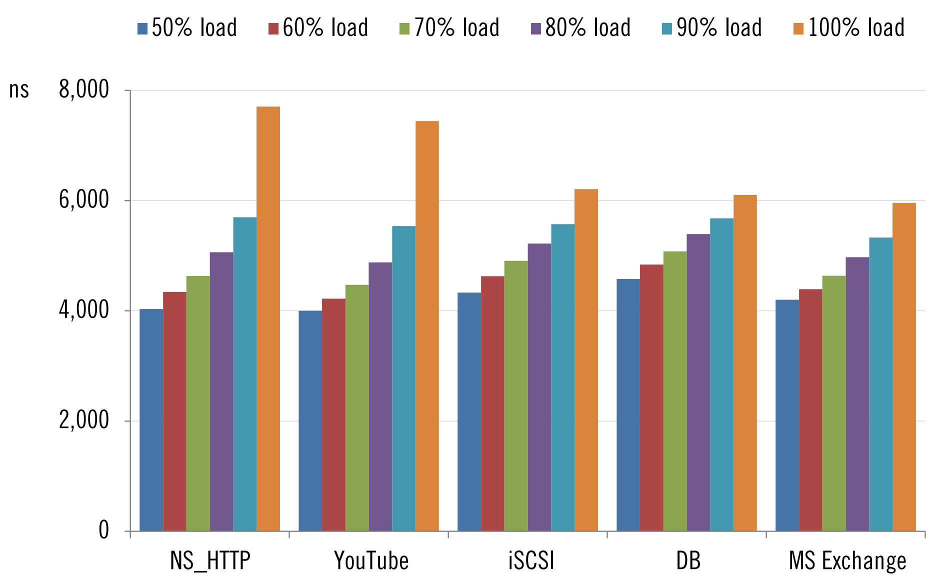
Arista 7500E RFC 3918 IP Multicast Test

288-40GbE or 1152-10GbE Accumulated Multicast Test Mode



	64	128	256	384	512	1024	1280	1518	2176	9216
Agg Tput Rate (%)	80.21	100.00	91.96	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Agg. Average Latency	11201.69	3979.76	52952.25	3969.76	3961.54	4010.10	3994.78	3957.41	3915.80	4184.03

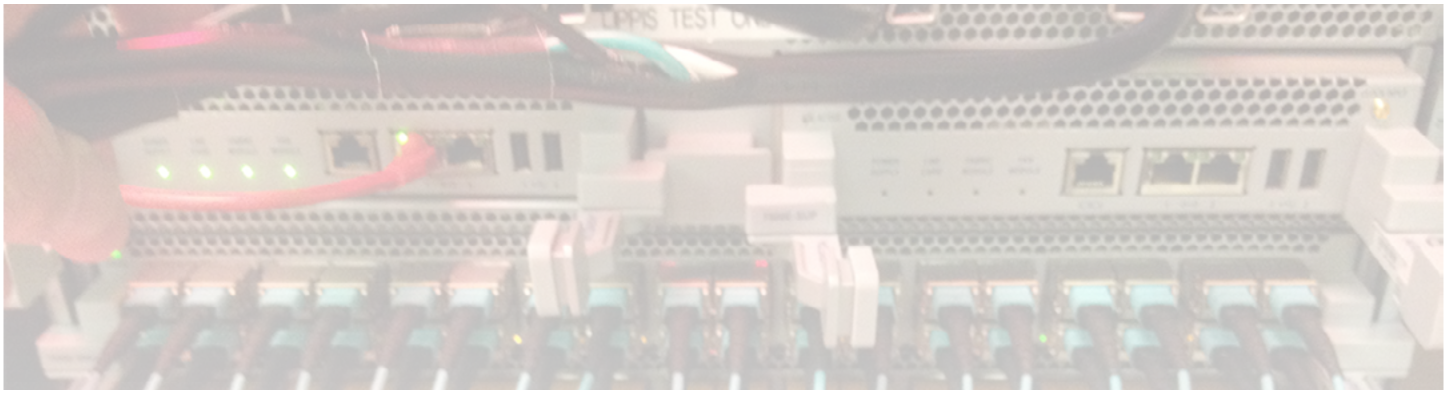
Arista 7500E
IxNetwork's CloudPerf Quick Test
(avg latency)



The Lippis defined CloudPerf test iMix was employed to generate traffic and measure latency plus throughput between ingress and egress. The Lippis defined CloudPerf test iMix consisted of east-to-west database, iSCSI and Microsoft Exchange traffic plus north-to-south HTTP and YouTube traffic. To understand Arista's 7500E performance under load, we ran six iterations of the Lippis defined CloudPerf test at traffic loads of 50%, 60%, 70%, 80%, 90% and 100%, measuring latency and throughput. All 288 40GbE ports were connected to Ixia test gear where IxNetwork's CloudPerf Quick Test iMix traffic was populated via 48 40GbE "north" and "south" ports plus 96 40GbE "east" and "west" ports. Therefore, 192 40GbE ports carried east/west flows while 96 40GbE ports carried north/south flows.

Average latency varied across protocol/traffic type as would be expected. Average latency within protocol/traffic type was stubbornly consistent as aggregate traffic load was increased, except for the 100% load iteration. At 100% load, some protocols experienced between approximately 35% more latency than at 90% load. Microsoft Exchange, iSCSI and database traffic were the traffic types with near consistent and low latency as load increased. HTTP and YouTube traffic experienced longer time to process, yet between 50% to 90% load were surprisingly consistent. The difference in latency measurements between 50% and 100% of load across protocols was 3.7 microseconds, 3.4 microseconds, 1.9 microseconds, 1.5 microseconds and 1.7 microseconds, respectively, for HTTP, YouTube, iSCSI, Database and Microsoft Exchange.

The 7500E performed flawlessly over the six Lippis defined CloudPerf test iterations. Not a single packet was dropped as the mix of east-west and north-south traffic increased in load from 50% to 100% of link capacity.



With the Arista 7500E connected to Ixia test gear and populated with 288 40GbE ports, the engineering team completed the standard Lippis/Ixia iMix tests. This test consisted of distributing traffic load and measuring power consumption at 0, 30 and 100% of line rate. With power levels in hand, engineers then calculated the power consumption of the Arista 7500E on a 10GbE and 40GbE per port basis. To compare different electronic equipment, we then calculated the Arista 7500E's TEER (Telecommunications Energy Efficiency Ratio) value. TEER is a measure of network-element efficiency quantifying a network component's ratio of "work performed" to energy consumed. Note that a larger TEER value is better as it represents more work done at less energy consumption.

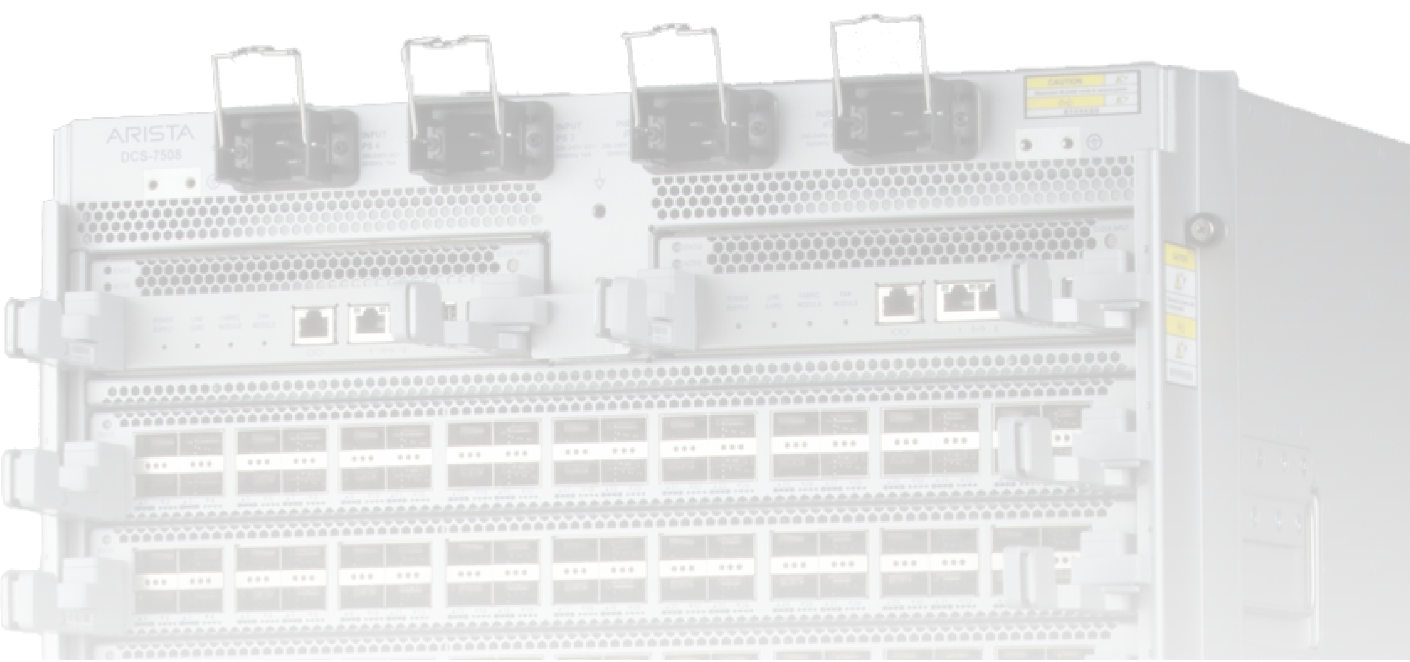
We found that the Arista 7500E is one of the most power-efficient modular data center switches in the industry with a power draw per 10GbE port of only 4.33 Watts and 17.30 Watts per 40GbE port, inclusive of optics and live traffic. The three-year cost to power the

Arista 7500E Power Consumption Test

Watts_{ATIS}/10GbE port	4.33
3-Year Cost/Watts_{ATIS}/10GbE	$15.81
1-Year Cost per 10GbE	$5.27
Total power cost/3-Year	$18,217.95
3 yr energy cost as a % of list price	1.42%
TEER Value	219
Estimated Annual power cost	$6,072.65
Cooling	Front to Back

Arista 7500E is estimated at $18,217.95 when fully populated with 1,152 10GbE ports; annualized, the 7500E is estimated to cost some $6,072.65 to power. The Arista 7500E's TEER value is 219. Previous data center modular switch TEER values ranged from a high of 246 to a low of 44, making the Arista 7500E one of the most power-efficient modular switches on the market.

The power consumption test demonstrated that the Arista 7500E achieves low power consumption due to efficient design that incorporates a compact system with front-to-rear airflow and an N+1 redundant fabric with efficient cooling fans that delivers a combination of low power draw, high density and support for triple speed 10/40/100GbE. The Arista engineers lead the industry in designing systems with deep buffers and a VOQ architecture that does not compromise power efficiency while delivering performance and redundancy.



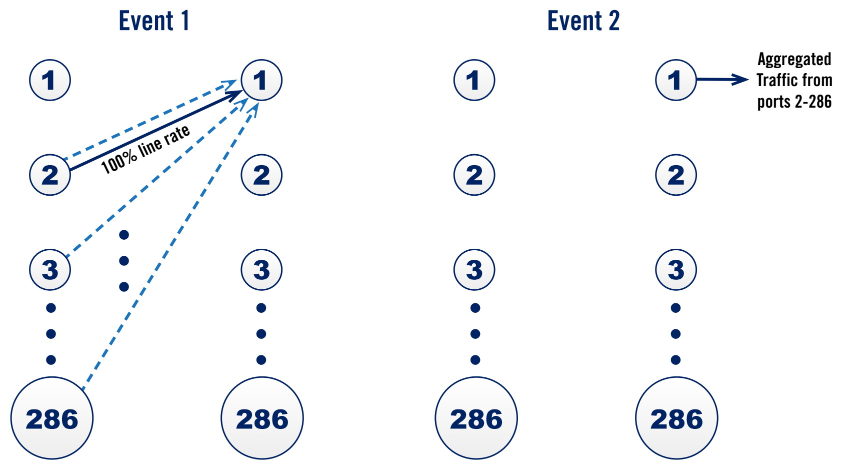
Additional Tests

The above tests are the standard Lippis Test conducted with Ixia since 2010. To verify Arista's claims that the 7500E provides adjustable VOQ buffer allocation, high performance 100GbE and 64-way ECMP evenly distributed hashing, we conducted and report on the following tests.

Adjustable VOQ Buffer Depth Test: To verify that VOQ buffer allocation is adjustable, two test cases were created: a small and large buffer test. In both small and large buffer tests, the configuration was the same; that is, a port pair (port 1 and 2) was supplied with unidirectional line rate or wirespeed traffic flow. Simultaneously, all ports (286) except port 2 transmitted a burst of traffic to port 1; that is, many ports transmitted to port 1. There were two events per test. Event 1 was the simultaneous transmitting of traffic from port 2-to-port 1 and many-to-port 1. Event 2 was the halting of unidirectional traffic between port 2 and port 1 and measurement of traffic flowing out of port 1, which should be equivalent to the many-to-port 1 aggregated traffic. The small buffer test used the 1518 byte size packet to transmit 34MB of load from all ports, except port 2 to port 1, or the many-to-port 1 load. The large buffer test used the 2000 byte size packet to transmit 373MB of load from all ports, except port 2 to port 1, or the many-to-port 1 load.

The goal of these two tests was to verify that Arista's 7500E VOQ buffers are adjustable and thus absorb traffic above wirespeed of 33MB and 373MB per port, respectively, without loss.

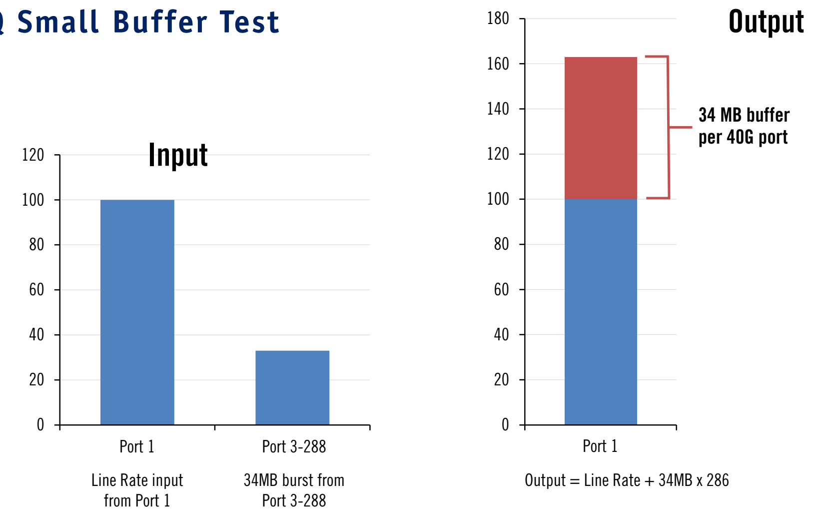
Adjustable VOQ Buffer Test



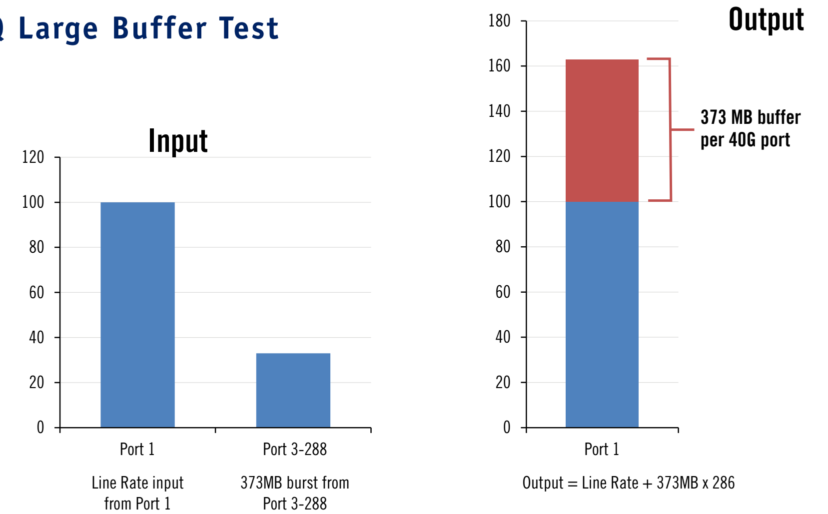
Test 1: Small buffer = many-to-port 1 34MB per port @ 1518 packet size

Test 2: Large buffer - many-to-port 1 373MB per port @ 2000 packet size

VOQ Small Buffer Test

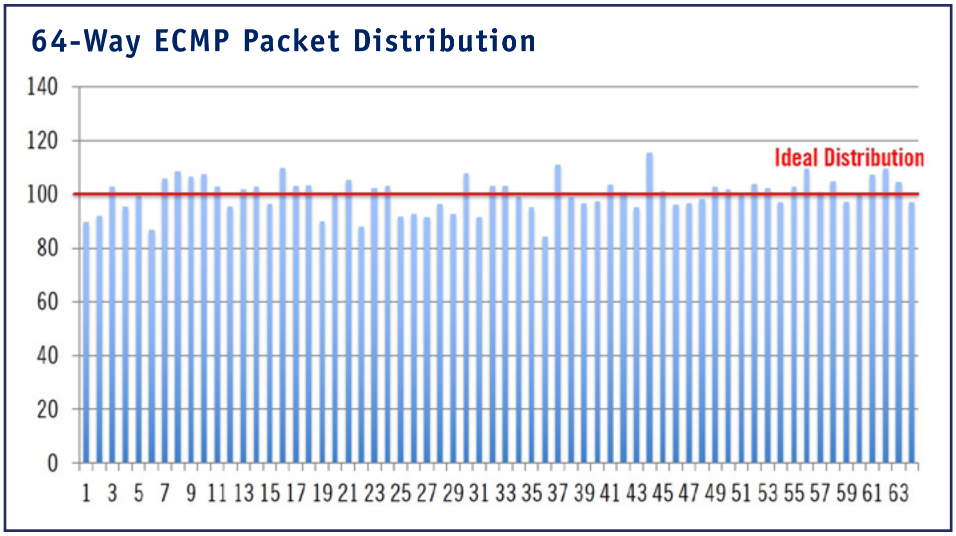


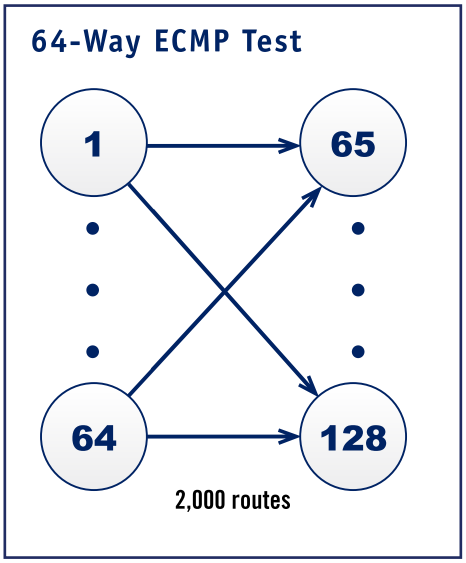
VOQ Large Buffer Test



We observed zero packet loss and expected average latency per port in the many-to-port 1 event 2 scenarios. Thus the Arista 7500's VOQ did adjust its buffer allocations that are resident within line cards and addresses misconceptions about "buffer bloat" that has been discussed in the industry.

64-Way ECMP Test: To verify that Arista 7500E's configured in a large 64-way two-tier network design evenly distributes traffic across spine switches, we configured a 7500E with 64 40GbE ports and tested its hashing algorithm. This was a simple test where we measured packet distribution across all 64 ECMP links. The configuration consisting of 20,000 routes was injected as 64-way ECMP through interfaces 65 to 128. Traffic from ports 1 through 64 was sent to the 20,000 routes at 50% of line rate. The graphic displays a near ideal distribution of traffic across all 64 ECMP links.





A Word about High Availability

The 64-way ECMP configuration is an excellent example of high reliability. High reliability is achieved through multiple design decisions. For a single switch, high reliability can be achieved with a low number of components—that's chips, fans, etc. Software reliability is the largest contributor to system failure and long mean time between failures (MTBF). Arista approached its EOS software architecture differently than most with a compartmentalization approach so that one application or network service does not bring down the entire operating system. Then there is network design; that is, how much redundancy is built between servers and leaf switches and between leaf and spine switches? It's the combination of all three designs that deliver high availability.

7500E Line Card Design and MTBF

High availability can be modeled and simulated to estimate the MTBF of the hardware, software and network system. But these estimates seldom match up with actuals.

The 7500E offers four line card choices:

- 36x40G, wirespeed QSFP ports and 18GB packet memory (wirespeed on all ports above 384 bytes average packet size)
- 12x100G, wirespeed, with integrated optics and 18GB packet memory. Can be configured as 144x10G, or 36x40G, or 12x100G via user configuration or mix-and-match
- 48xSFP+ and 2x100G uplinks with built in optics and 9GB packet memory
- 48xSFP+ and 9GB packet memory

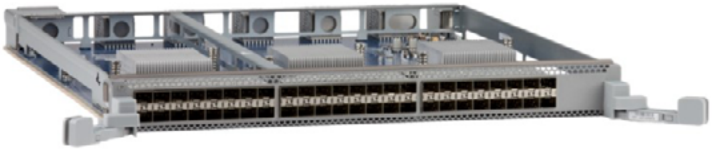
12x100GbE

36x40GbE

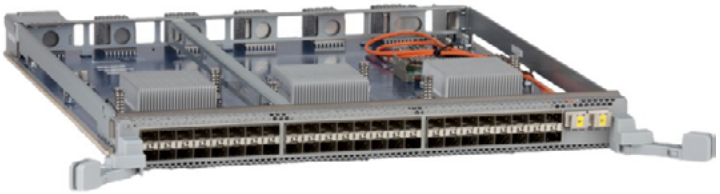
The 36x40G line card is the most complex of these and has six packet processors. These packet processors have network ports wired through the QSFP connectors on the front and fabric interfaces going through the back. 3GB of packet memory sits next to each packet processor. These line cards are designed using high-speed signaling, with very high-quality, high-layer count PCB (Printed Circuit Board). While there are several components on the board, most of these are "decoupling" capacitors, and there are so many that a few capacitors, or resistors, failing over the life of the product have no impact to functionality or performance. The reliability of the product is tied to the most critical components on the board, and these products have high-calculated MTBF based on the latest Telcordia standards (high MTBF is good).

The Arista 7500E has been in production since June 2013, and based upon company data, its observed MTBF in the field is in the 250,000- to 290,000-hour range for the 36x40G line cards and even higher for the other models. Hence network operations teams can be confident with the design and reliability of the product.

48x10GbE



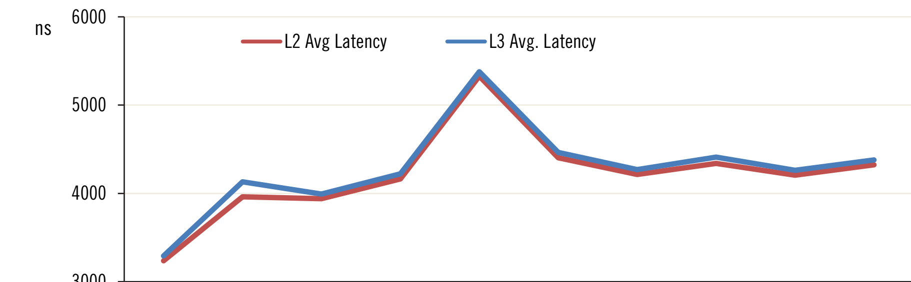
48x10GbE + 2x100GbE



100GbE Line Card Performance Test: To verify line rate throughput performance of the DCS-7500E-12CM-LC line card with embedded optics, RFC 2544 latency and throughput tests were conducted. The DCS-7500E-12CM-LC line card was configured with two 10GbE and two 100GbE ports connected to Ixia test gear to stress the 100GbE capabilities of the 7500E.

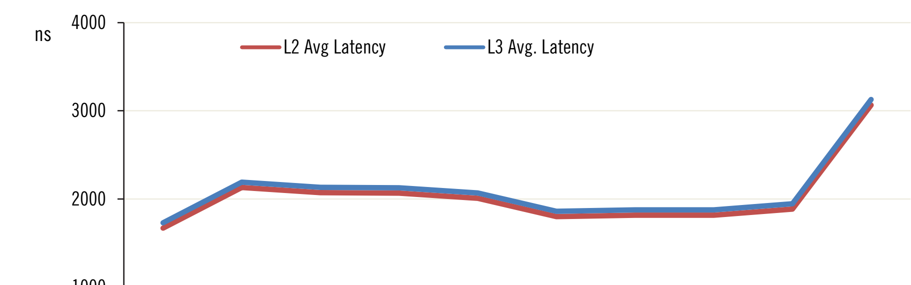
In the 100/10GbE test configuration, the 7500E produced average latency that ranged from a low of 3.3 microseconds (3277 ns) to a high of 4.4 microseconds (4364.5 ns) for L2 traffic forwarding across 64 and 9216 byte size packets measured in store-and-forward mode at 100GbE at 100% line rate. The 10GbE ports demonstrated average latency that ranged from a low of 1.7 microseconds (1669.5 ns) to a high of 3.1 microseconds (3066.5 ns) for L2 traffic forwarding across 64 and 9216 byte size packets measured in store-and-forward mode.

Arista 7500E RFC2544 Layer 2 & Layer 3 Latency Test
2-100GbE



	64	128	256	384	512	1024	1280	1518	2176	9216
L2 Avg Latency	3277.0	4000.5	3981.0	4203.5	5369.0	4443.0	4255.0	4380.0	4245.5	4364.5
L3 Avg. Latency	3270.5	4111.0	3972.0	4202.0	5358.0	4445.0	4249.5	4391.0	4240.5	4358.0

Arista 7500E RFC2544 Layer 2 & Layer 3 Latency Test
2-10GbE



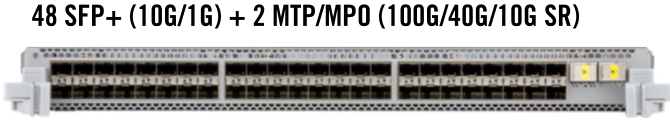
	64	128	256	384	512	1024	1280	1518	2176	9216
L2 Avg Latency	1669.5	2128.5	2070.0	2064.5	2006.0	1797.5	1815.0	1816.0	1885.0	3066.5
L3 Avg. Latency	1669.0	2128.5	2070.0	2064.5	2005.5	1797.5	1815.0	1816.0	1884.5	3067.5

For L3 forwarding, the 100GbE test produced average latency that ranged from a low of 3.3 microseconds (3270.5 ns) to a high of 4.4 microseconds (4358 ns) across 64 and 9216 byte size packets measured in store-and-forward mode. The 10GbE ports demonstrated average latency that ranged from a low of low of 1.7 microseconds (1669 ns) to a high of 3.1 microseconds (3067.5 ns) across 64 and 9216 byte size packets measured in store-and-forward mode. Note the latency measurements for L2 and L3 forwarding was, for all practical purposes, identical across the 10GbE and 100GbE test data. Also note that throughput was 100% for both 100GbE and 10GbE across all packet sizes.

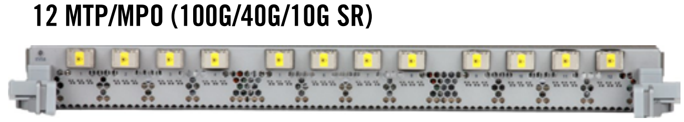
36 QSFP+ (40G/10G)

48 SFP+ (10G/1G)

48 SFP+ (10G/1G) + 2 MTP/MPO (100G/40G/10G SR)



12 MTP/MPO (100G/40G/10G SR)



Discussion

The Arista 7500E is the only spine switch that offers both L2 and L3 forwarding at 288 40GbE scale and at ultra low latency. It's also the only modular switch we have tested that offers 10/40 and 100GbE line card options. It offers the best congestion management system measured to date, thanks to its generous buffer allocation and VOQ buffer algorithm. Also impressive is that the 7500E's buffers can be adjusted with a single command. At 4.33 Watts per 10GbE, it is one of the most power efficient modular switches available.

The Arista 7500E was designed to solve two massively disruptive IT trends: cloud and mobile computing. Distributing computing tools or the ability for a small number of operational personnel to manage hundreds to thousands of servers is one of the major drivers of cloud computing economics being leveraged by cloud providers and enterprise IT. In fact, the number of servers per rack has grown from 40 in 2010 to 80 in 2013—a trend that will only continue along with the increased number of servers being connected into the cloud network at 10GbE. Along with server density increasing, application access via mobile devices is driving massive east-west and north-south data/traffic flows, thanks to content being distributed across many servers. From a networking perspective, all this adds up to increasing number of 10GbE server connections via leaf or Top-of-Rack (ToR) switches and subsequent 40GbE connections between leaf and spine switches to provide high performance and low latency "Brownian motion" traffic flows. These high density and performance plus low latency demands are signaling the need for a new generation of 10/40/100GbE modular switching in private and public cloud infrastructure plus high-end data centers.

The Arista 7500E was engineered for these trends, and it demonstrated its engineering achievements at this Lippis/Ixia test. The 7500E delivered one of the lowest latencies observed for modular switching over the past three years at the Lippis/Ixia industry test and at the top of the range of 40GbE port densities we have tested. The 7500E is only one of two modular switches tested for performance while processing 288 ports of 40GbE or 11.5 Tbps of wirespeed traffic.

For IP multicast traffic forwarding in the distributed mode, the Arista 7500E demonstrated ~0.35% aggregated throughput with no packet loss plus store-and-forward latencies ranging from a low of 3 microseconds (3002.3 ns) at 64 byte size packets to a high of 3.5 microseconds (3547.1 ns) at 9216 byte size packets. The Arista 7500E's congestion management was unparalleled at 77% to 82% of aggregated forwarding rate as percentage of line rate during congestion conditions for L2 and L3 traffic flows. Considering the density of 40GbE ports supported and sheer magnitude of the traffic flowing into the 7500E, the engineers achieved congestion management at a scale never before observed.

Based upon these tests, the Arista 7500E can enhance the performance of various data center applications. In a big data network architecture scenario, the 64-way ECMP configuration offers a two-tier network architecture that can scale to tens of thousands of servers. This same two-tier cloud network architecture with latency between 1.6 and 4 microseconds that's consistent to within a few nanoseconds is well engineered to support IP storage. Arista has announced a wide variety of interoperable storage partners and architectures with solutions from Isilon, Nutanix, NetApp, EMC, Panasas, SolidFire, DAS, NAS, iSCSI, Lustre, etc., that should find the 7500E well suited to support their IP storage reliably and at speed.

The Arista 7500E was designed to enable a Universal Cloud Network that's software provisioned and supports any application, anywhere in the network. The 7500E boasts advanced traffic control and monitoring with high density and scalable network resources. It's built for highly virtualized environments—thanks to its integrated wirespeed network virtualization support—is power efficient with rich software resources and engineered with deep buffers to mitigate congestion and increase reliability.

The Arista 7500E became generally available in quarter 2 of 2013. With pricing of 10GbE at $550 per port, 40GbE at $2,000 per port and 100GbE at $8,500 per port, including optics, the 7500E sets the new standard for high-density modular switch performance and economics. For L2 and L3 forwarding, 40GbE aggregation, hyperscale and spine switch use cases, the Arista 7500E proved during these Lippis/Ixia tests that it's well engineered for these environments.

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.