

25G Ethernet: A new standard for network connectivity

Cost and performance optimized speed for server to switch connections



The undeniable growth in demand for bandwidth is driving data center networks for better scalability, instant bandwidth provisioning and application agility. Data centers are expanding at an unprecedented pace as is the virtualization density, driving the need for higher bandwidth between the server and switches. Multi-core processing power, NVMe and PCIe 3.0 I/O buses, and high-speed flash storage require networking that can keep pace with increased throughput.

To meet these needs, networking and the Ethernet industry are moving in a new direction. In addition to 10, 40 and 100GbE networking, 25G and 50G Ethernet technologies are the new standards that offer significant density, cost and power benefits for server to top of rack connections. This paper explains why and shows how changes in demand, technology and network economics are working together to alter the networking landscape and how Arista and Cavium are working together, spearheading the effort to bring solutions to the market.

Networking at the crossroads

As organizations push more traffic into their enterprise and private data centers, and with bigger “pipes” coming into the enterprise, more bandwidth is required as data moves down into the fabric or top of rack. ToR architectures utilizing 10GbE to the server are among the largest number of connections in data centers today. But these architectures are rapidly outgrowing 10GbE capabilities, often requiring multiple ports.

40GbE technology has evolved over the years and has gained some momentum as an option for enterprises, service providers and cloud providers looking to scale beyond 10GbE. The 40GbE ecosystem is complete with switching and NIC solutions from multiple vendors including Arista and Cavium/QLogic. However, since the underlying technology for 40G Ethernet is simply four lanes at 10G speed, it does not offer the cost per bit, power consumption or server rack density advantages which are the necessary enablers for widespread speed transition.

As the Ethernet industry continues to innovate and lay a path to higher networking speeds like 100GbE, 25GbE has been developed to provide a simpler path to future Ethernet speeds of 50 Gbps, 100 Gbps and beyond. These implementations capitalize on the 25GbE specification adopted by the 25 Gigabit Ethernet Consortium. The specification makes use of single-lane 25 Gbps Ethernet links and is based on the IEEE 100GbE standard (802.3bj).

With the release of the 25GbE specification, new generations of both switching silicon and adapter chips are now available. Over the last few years, the underlying high-speed signalling technology has advanced to the point that 25 GHz channels and interfaces are economically viable, and the physics-related challenges in signal integrity have found reliable solutions.

The need for interoperability

As with any new technology, one of the biggest challenges is the uncertainty of how different systems in the network communicate effectively and consistently. While the standardization effort by IEEE alleviates some of this concern, there are still several pieces to the puzzle that can be solved with interoperability testing between Ethernet switches and Network interface cards. Auto-negotiation of speed, forward error correction support and the loss characteristics of the physical media device that connects the systems are some of the critical parameters to ensure error-free communication in the network. There is an ongoing interoperability effort in the industry that is being lead by Ethernet Alliance and University of New Hampshire interoperability lab but Arista and Cavium are going beyond this to ensure seamless customer deployments of 25G Ethernet.

Partnering for your success

As members of the 25 Gigabit Ethernet Consortium, Cavium/QLogic and Arista are dedicated to bringing the benefits of the new technology to their customers. Arista is one of the first switching companies to release a wide portfolio of IEEE compliant 25G and 100G switches in fixed-form factor as well as chassis based solutions for leaf and spine deployments. Cavium/QLogic was the first to demonstrate an end-to-end 25GbE solution. Both Arista and Cavium 25GbE technologies are backward compatible with 10GbE, allowing customers to build and cross-connect a heterogeneous-speed Ethernet network.

Arista and Cavium have partnered to test Switch and NIC interoperability with Arista’s 7160 and 7060X2 platforms with Cavium’s FastLinQ QL41 1x series of NIC cards. Later part of this whitepaper describes the test setup, test cases and test results.

The economics of 25GbE

The availability of 25GbE capable switches and adapters has dramatically affected the economics of server-to-switch connectivity upgrades. Upgrading from 10GbE to 25GbE speeds offers both CapEx and OpEx savings through backward compatibility, for investment protection and seamless migrations with consistent rack-design and reuse of the existing cabling infrastructure, avoiding costly and complex changes.

Recent high performance 100/25GbE chips use single-lane 25G serializer-deserializer (Serdes) technology similar in operation to 10GbE but delivering 2.5 times the performance, thus reducing the power and cost per gigabit significantly. This also translates to higher system bandwidth at the same power consumption compared to equivalent 10G solutions. In addition, 25G provides higher port and system density than a comparable 40G solution (Table 1). Both power savings and higher density results in lower cooling requirements and operational expenditure for data center operators.

Table 1: Power and Cost savings of 25G/100G solution over 10G/40G solution

| Top of Rack Switch | 48 x 10G | 48 x 25G | 32 x 40G | 32 x 100G |
|-------------------------------------|----------|----------|-----------|-----------|
| Servers | 48 x 10G | 48 x 25G | 128 x 10G | 128 x 25G |
| Total Bandwidth | 0.4 Tbps | 1.2 Tbps | 1.28 Tbps | 3.2 Tbps |
| Power/Gigabit: Switch + Cables | X | 0.5X | 0.4X | 0.25X |
| Cost/Gigabit: Switch + NIC + Cables | Y | 0.4Y | 0.5Y | 0.25Y |

As the majority of 25G switches and Network Interface cards offer backward compatibility to 10G, there is the flexibility to manage a gradual migration to higher speed servers and mix and match port speeds. All SFP based 25G ports on Arista switches and 25G NICs from Cavium can be used at 10G speed. With a minimal premium for 25G based systems compared to 10G systems, it becomes a simple decision to deploy 25G capable systems, operate today at lower speeds and migrate to 25G over time, on a per server basis, and realize the performance advantages.

As servers and switches utilize the SFP form factor for 25G optics and cables, they leverage the smallest and most widely deployed form factor for 1G and 10G, with the same high port density and familiar layout, allowing the benefit of reusing the same server rack designs and working with familiar cabling infrastructure. The rack configuration, layout and cabling infrastructure can remain exactly the same. For example, if the current design is a 32x40G leaf switch with 10G servers connected via 40G to 4x10G break-out direct-attach cables (DAC), the same rack density and design can be realized with new 32x100G leaf switches. New 25G servers are connected to the switch with 100G to 4x25G break-out DACs. Similarly, if the leaf is currently a high-density 10G SFP switch with multiple QSFP 40G uplinks to the spine, the same server port density is available, with SFP for 25G servers and QSFP for 100G uplinks. This allows for a simple and seamless migration.

Considering the significant benefits and compelling economics of 25G, it is no surprise that the move to 25GbE (and 50GbE) is accelerating — a recent five-year forecast by industry analysts at the Dell'Oro Group predicts 25GbE will be the dominant Server port speed for new systems by 2018.²

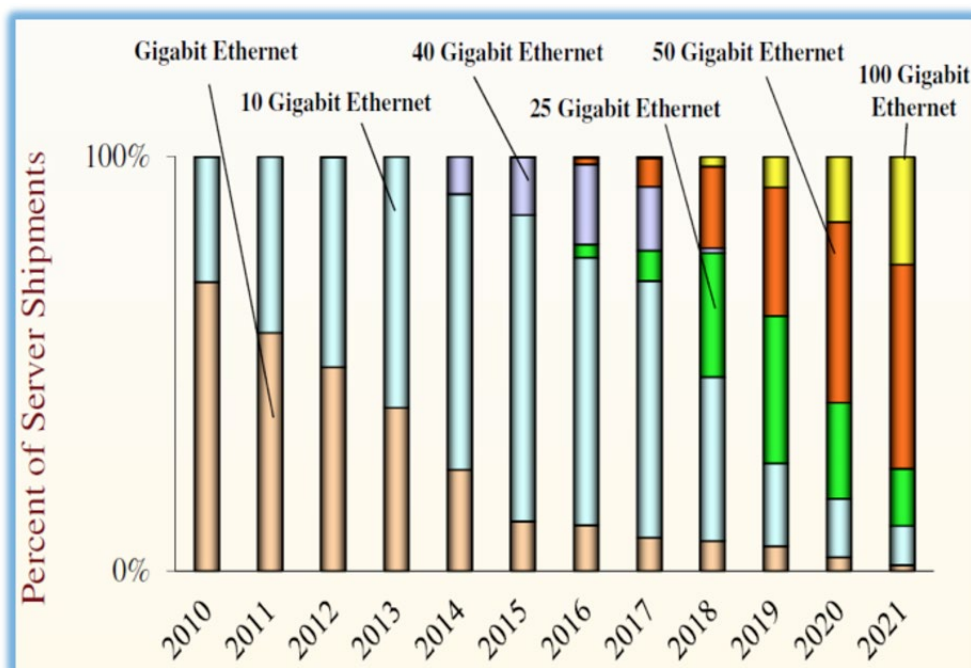


Figure 1: Server Adapter Port Adoption in Cloud. Source: Dell'Oro Research, 2017

Key factors enabling 25GbE advances and economics include:

New silicon

- 25G SerDes based silicon development now allows high-density 100GbE switches
- 32x100GbE or more switching bandwidth is now available on a single chip
- High performance switches offer the capability to run 4x25GbE from a 100GbE port

New optics

- 25G optics/cables introduced with smallest size offering investment protection
- Familiar fiber and cabling infrastructure with single lane 25G and 4 lane 100G
- Lower power and cost/bit compared to 10G

New economics

- 25GbE helps reduce CapEx with fewer ToR switches for a given bandwidth requirement
- It also reduces OpEx in terms of power, cooling and space
- 25 Gbps Ethernet adapters adherent to IEEE standards are now available, helping protect technology investments

25GbE Standards

Broad Industry Support

The new 25GbE specification was championed by the [25 Gigabit Ethernet Consortium](#), which consists of leading companies including Arista Networks™, Broadcom®, Brocade®, Cavium, Cisco®, Google®, Microsoft® and many others. The goal of the consortium was to promote the standardization and use of 25 and 50Gb an industry-standard, interoperable Ethernet specification that boosts performance and slashes interconnect costs in order to benefit consumers and the industry.

IEEE Standardization Status

In July 2014, the IEEE held a “call for interest” meeting and members unanimously agreed to support the development of a 25GbE standard for servers and switching. The IEEE P802.3by 25Gbps Ethernet Task Force was formed to develop the standard. In addition, the IEEE P802.3bq 40GBASE-T Task Force adopted objectives to develop BASE-T support for 25GbE. Both of these task forces have already completed the technical work and published the standards. The fast progress was due to the high degree of leverage from existing standards for 100GbE as the base for 802.3by and 40Gbps twisted pair development for the 25GBASE-T specifications.

The IEEE 802.3by specification contains 3 different 25G cable types (CA-N, CA-S, CA-L), which are aligned to 3 different loss classification categories and associated minimum FEC requirements.

Table 2: IEEE defined Direct Attach Cable Type

| Cable Type | Loss (dB) | Notes |
|------------|-----------|--|
| CA-N | 12.98 | Highest grade, No FEC required |
| CA-S | 16.48 | Medium grade, Firecode/FC-FEC required |
| CA-L | 22.48 | Lowest grade, RS-FEC required |

Arista leading with standards compliant 25G Portfolio

Arista Networks, as one of the original founders of the 25G Ethernet consortium, has been a champion of 25G and 50G Ethernet standardization. As a leader in high-speed Ethernet switches that redefined networking architectures, Arista envisioned the significance of 25G and 50G Ethernet speeds and their dramatic price and performance impact on modern data center networks.

Arista is focused on building a range of standards compliant 10/25/50/100 Gigabit Ethernet (GbE) switches that continue to define next generation network architectures, bring extensibility to networking and dramatically change the performance of data center networks. Arista Network's range of leaf and spine switches include systems with support for 100GbE using QSFP ports that can be converted to 4 lanes of 25GbE and a growing range of systems with support for IEEE compliant 25GbE ports using the SFP form factor. Each system that supports 100G or 25G mode can support 40G or 10G, respectively, for deployment flexibility and cost optimization in the network. It is the multi-speed capability that provides the necessary backward compatibility that allows for seamless and cost effective upgrade paths as well as the ability to mix and match interface speeds during upgrade and migration cycles. The Arista 7060X, 7160 and 7280R Series are the first systems to offer full support for the IEEE 25GbE 802.3by standard. Future 100G and 25G systems will continue to expand these choices.

The Arista 7060X and 7260X Series are a range of 1RU and 2RU high performance and density 40GbE and 100GbE fixed configuration data center switches with wire speed layer 2 and layer 3 features, combined with advanced features for software driven cloud networking. The 7060CX2-32S offers the flexibility to break-out to 128x25G ports with full compliance to IEEE 802.3by specification.

The Arista 7160 Series can be deployed in a wide range of open networking solutions including large scale layer 2 and layer 3 cloud designs, overlay networks, and virtualized or traditional enterprise data center networks. The 7160-48YC6 offers 48 native 25G ports for server connectivity and six 100G ports for uplinks. The 7160-32CQ is a high density 100G switch that can be used in many flexible port combinations including breaking out of every 100G port to 4x25G for up to 128x25G ports in a single system.

The 7280R series are a set of purpose built 10/25/40/50/100GbE fixed configuration 1RU and 2RU systems designed for the highest performance environments such as IP Storage, Content Delivery Networks, Data Center Interconnect and IP Peering. The 7280SR2-48YC6 offers 48 native 25G ports along with six 100G uplink ports. Other 7280R Series systems offer 4x25G support on QSFP 100G ports with the capability to mix and match 10/25/40/50/100G speeds.



Figure 2: Arista Networks standards compliant 25G Switch portfolio

Cavium FastLinQ QL41000 Series

25 Gbps Ethernet-to-PCIe 3.0 Intelligent Ethernet Adapters

Cavium delivers the industry's most comprehensive network adapter portfolio – FastLinQ® Standard Ethernet Adapters, Converged Networking Adapters and LiquidIO™ Intelligent NICs that cover the entire spectrum of customer Ethernet connectivity and offload requirements

Purpose built for accelerating and simplifying data center networking, Cavium FastLinQ Ethernet technology delivers:

Broad Spectrum of Ethernet Connectivity Speeds – 10/25/40/50/100GbE to host the most demanding enterprise, telco and cloud applications to deliver scalability to drive business growth.

Enterprise Class Reliability and Flexibility – With millions of Ethernet ports shipped, and a flexible architecture that delivers faster time to market and adaption to new and emerging technologies, FastLinQ Ethernet NICs are amongst the top choices for enterprise data centers.

Cloud Ready – Industry's only network adapter with Universal RDMA, that delivers the ultimate choice and investment protection with concurrent support for RoCE, RoCEv2 and iWARP. Acceleration for Network Virtualization by offloading protocol processing for VxLAN, NVGRE, GRE and GENEVE enables customers to build and scale virtualized networks without impacting network performance.

Storage Acceleration - Full protocol concurrent offload for NVMe-oF™, iSER, iSCSI and FCoE delivers up to 5M IOPS while consuming the fewest server CPU cycles, leaving headroom for virtual applications and higher RoI on server investments with seamless upgrade paths of next-gen storage connectivity.

Telco and Network Function Virtualization (NFV) – Leading small packet performance and integration with DPDK and OpenStack, enables Telco's and NFV application vendors to seamlessly deploy, manage and accelerate the most demanding NFV workloads.

Cavium FastLinQ 41000 Series

2nd Generation Feature-rich 10/25/50Gbps Performance NIC

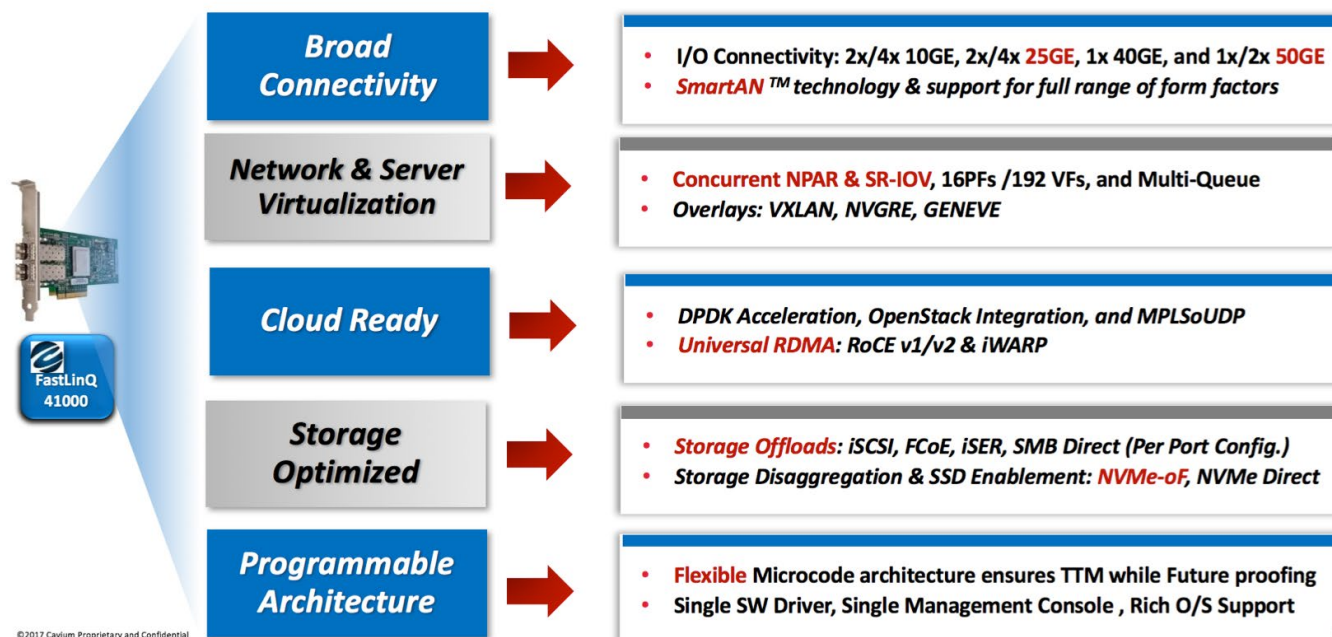


Figure 3: Cavium FastLinQ 41000 Series 10/25Gbps Ethernet NIC with Universal RDMA

With superior price/performance versus 10GbE, the powerful Cavium FastLinQ QL41000 Series 25Gbps Ethernet Adapters available from leading OEMs and ODMs in standard, mezzanine and OCP form factors are optimized for use across enterprise, managed service provider (MSP) and large public cloud deployments.

Interoperability Test Set-up

The Arista Switch is connected to two servers, each with Cavium 25G NIC cards. An iPERF server is used to generate traffic on one server, and data is transmitted through the Arista switch to the second server with an iPERF client. For connectivity, 3 and 5 meter SFP-SFP 25G Twinax copper cables and QSFP-4xSFP breakout cables were used (Figure 4).



Figure 4: Test set-up for 25G Ethernet Switch-NIC interoperability testing

Table 3: Test Cases

| No. | Test |
|-----|---|
| 1 | Link bring up between the Arista Switches & NIC cards |
| 2 | Link shut/no shut from both ends of connection |
| 3 | Verify the interop in Auto-negotiation mode |
| 4 | Verify the interop for different speeds |
| 5 | Link pull/push from both ends of connection |
| 6 | Reload the DUT / Server and check the link status between DUT & NIC |
| 7 | Configure IP addresses on both ends, verify line rate traffic using iperf |

Table 4: Test Results

Arista Switch platform: 7160 - 48YC6, Cavium 25G NIC: QL41262HLCU

| Test Mode | 3m SFP-SFP 25G Cable 26AWG, CA-N ¹ spec | 5m SFP-SFP 25G Cable 26AWG, CA-L spec |
|---|---|--|
| Speed: 25G Forced FEC ² : RS-FEC Forced | Pass | Pass |
| Speed: 25G Forced FEC: FC-FEC Forced | Pass | Pass |
| Speed: 25G Forced FEC: Disabled | Pass | N/A (FEC required) |

Table 5: Test Results**Arista Switch platform: 7060CX2-32S, Cavium 25G NIC: QL41262HLCU**

| Test Mode | 3m Q-S 100G-25G Cable 26AWG, CA-S spec | 5m Q-S 100G-25G Cable 26AWG, CA-L spec |
|---|---|---|
| Speed: 25G Forced FEC ³ : RS-FEC Forced | Pass | Pass |
| Speed: 25G Forced FEC: FC-FEC Forced | Pass | N/A (RS-FEC required) |
| Speed: 25G Forced FEC: Disabled | N/A (FEC required) | N/A (RS-FEC required) |

Conclusion

The cost, power and performance benefits of 25G Ethernet technology are undeniable. Backward compatibility for staged migration, familiar cabling infrastructure and investment protection with standards compliance are strong reasons for data center operators to migrate to the next generation Ethernet speed. Arista Networks and Cavium, the industry leaders in high speed Ethernet switching and Network Interface cards, respectively, are key enablers of the ecosystem and have partnered to provide tested solutions which customers can deploy with confidence. The new standard for network connectivity has arrived!

¹ CA-N, CA-S and CA-L are the various cable types as explained in Table 2

² FEC: Forward Error Correction, RS-FEC: Reed Solomon FEC as defined by IEEE 802.3by, FC-FEC: Firecode FEC as defined by 25G Ethernet consortium

Santa Clara—Corporate Headquarters

5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500

Fax: +1-408-538-8920

Email: info@arista.com

Ireland—International Headquarters

3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office

9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

San Francisco—R&D and Sales Office

1390 Market Street, Suite 800
San Francisco, CA 94102

India—R&D Office

Global Tech Park, Tower A & B, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

Singapore—APAC Administrative Office

9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

Nashua—R&D Office

10 Tara Boulevard
Nashua, NH 03062

