Servers and storage rapidly adopting 25GbE and 100 GbE networking

A quiet revolutation is taking place in networking speeds for servers and storage, one that is converting 1Gb and 10 Gb connections to 25Gb, 50Gb and 100 Gb connections in order to support faster servers, faster storage, and more data. The switch to these new speeds is changing not just networking equipment but also network deployment models.

By John Kim, Chair SNIA Ethernet Storage Forum

The Need for Speed

Several technology changes are collectively driving the need for faster networking speeds.

· **Faster Servers** – The first driver of faster network speeds is the spread of faster servers. Modern x86 servers can push more than 100Gb/s of throughput, if their entire workload is focused on network I/O. While very few servers reach that level, it's increasingly common for one server to need more than the throughput that a 10GbE or 8Gb FC port can deliver. The recent robust competition between Intel, AMD, Open Power, and various ARM CPU vendors means server power will keep increasingly rapidly.

· **Denser compute** – As IT departments switch from virtualization to containers, the lower overhead of containers lets them pack more compute instances on each server than using virtual machines. This requires more storage performance and network throughput.

· **Faster Storage** – Everyone knows flash has been replacing disk, but in addition to the continuation of this trend, the flash drives now have faster NVMe interfaces, which deliver higher throughput and lower latency than SAS or SATA SSDs.

· **Bigger Data** – The amount of data and size of individual files and databases is growing rapidly due to increased capture of video, higher-resolution video (4K and 8K takes up approximately 4X and 8X more space and bandwidth than regular 1080p high-definition video), smart phones, social media, and the Internet of Things.

· **Distributed Compute** – Various applications distribute compute and storage across a cluster of servers. Big data (Hadoop and NoSQL databases), machine learning, scale-out storage, and hyperconverged infrastructure all do this and require high levels of "east-west" network traffic between the compute or storage nodes.

Networking Vendors Deliver

Any technology revolution requires both supply and demand. If customers want faster speeds but network vendors don't offer it—or if the prices are too high, then implementation will only be in niches and adoption will be slow. Fortunately the networking vendors are delivering 25/100GbE networking solutions to market quickly. There are an abundance of 25/100GbE NICs (network interface cards) and switches available today.

At least four switch silicon vendors (Broadcom, Cavium, Cisco, and Mellanox) sell 25/100GbE products, leading to a multitude of 25/100GbE switches from Cisco, Dell/EMC, HPE, Huawei, Lenovo, Mellanox, Quanta, Supermicro and others, including "whitebox" switches from original

design manufacturers (ODMs). Network adapter silicon and cards in these speeds are widely available from Broadcom, Cavium, Chelsio, Intel, Mellanox, Solarflare and others, with these cards being sold through many server original equipment manufacturers (OEMs), storage OEMs, resellers and systems integrators.

Competition and increasing volumes have already driven price premiums down so that 25GbE NICs cost only 20-30% more than 10GbE NICs, and 100GbE switches cost only 50% more per port than 40GbE switches. Customers enjoy 2.5 times more bandwidth than before, along with lower latency, at only slightly higher prices, and the trends show 25/100GbE network equipment pricing will move closer and closer to 10/40GbE pricing over time. As a bonus, existing fiber optic cabling designed for 10GbE can be used for 25GbE, and cabling designed for 40GbE can be reused for 100GbE, simply by installing new transceivers or modules at the ends of the cables.

Because servers are typically deployed for 2-4 years and switches for 3-5 years, it makes financial sense to “future-proof” networks by ensuring all new Ethernet adapters, switches and cables can support 25GbE (or 50GbE) endpoints and 100GbE links between switches.

Cloud Goes First, Technical Computing Second

The customers first to adopt these faster speeds have been hyperscalers (such as Facebook, Google, Microsoft or Amazon) and other large cloud service providers (SPs). They need the increased bandwidth and can easily change their server and network designs to take advantage of faster speeds and new features in the network. They are relentlessly driven to improve performance and efficiency, since pricing for cloud Infrastructure, Platforms, Software, and Storage-as-a-Service (IaaS, PaaS, SaaS, and StaaS) is constantly declining.

With large scale networks, the uplink speed between switches is a critical factor to maximizing efficiency and lowering costs. With 40GbE uplinks, each “leaf” or top-of-rack (ToR) switch might require 4-10 (160Gb/s to 400Gb/s) uplinks to the spine switches, which then require 4-10 uplinks to core routers. By using 100GbE the number of uplinks per switching layer can be reduced by 60%. For example, two or four 100GbE links would provide the same or greater bandwidth as using four or ten 40GbE links. The reduction in links allows for more servers per rack and reduces the floor space, power, and heat footprint of the network.

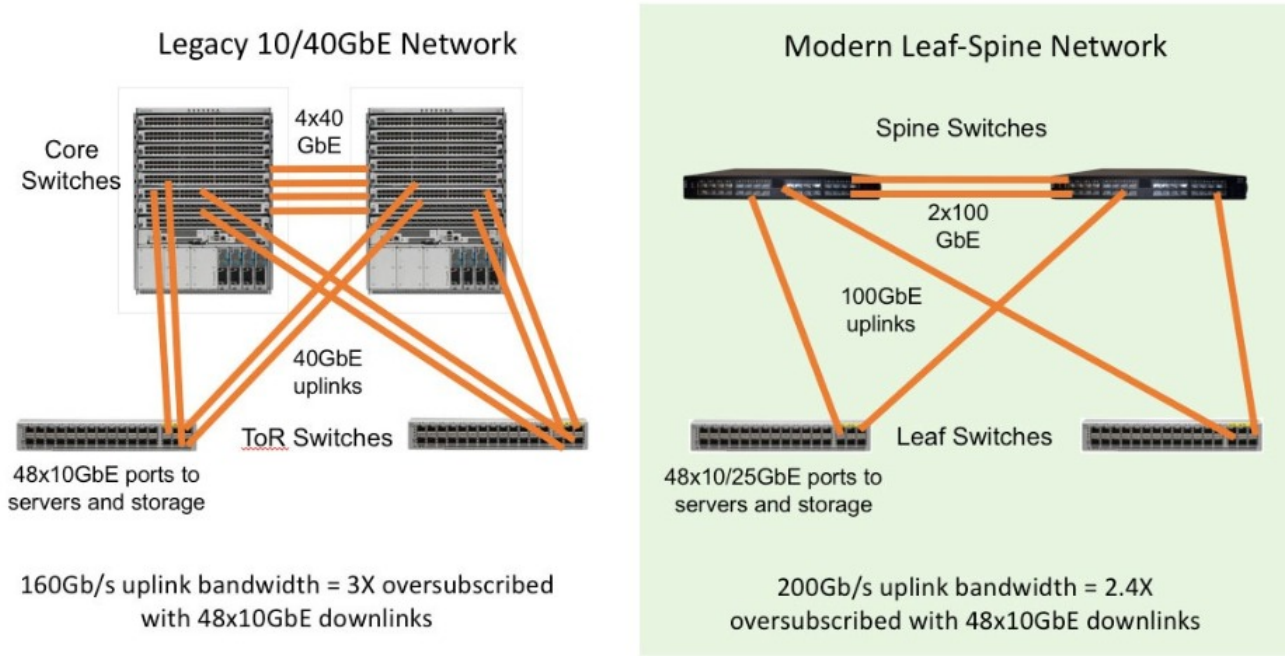


Figure 1: 100GbE uplinks allow for more bandwidth using fewer uplinks, simplifying the design of large networks.

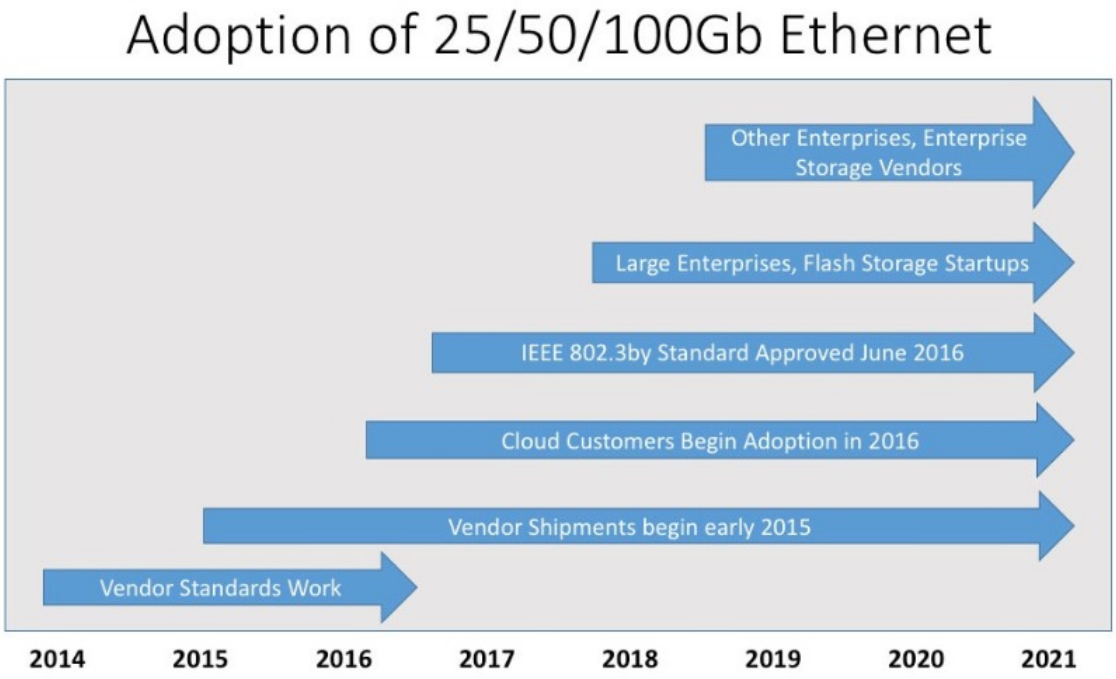
Because large cloud SPs deploy tens or hundreds of thousands of servers and hundreds of switches each year, they replace or install new racks, rows or even complete data centers all at once using newer and faster networking to support newer and faster servers and storage. The largest SPs have already moved new deployments to use 25 or 50GbE on the servers and storage, with 100GbE uplinks between switches and across the data center. Some high-end all-flash arrays with NVMe SSDs—sometimes with NVMe over Fabrics as the storage networking protocol—are also running 100GbE. Due to the desire for network uplink efficiency explained above, many of the largest cloud vendors are already demanding 200GbE and 400GbE connections between switches. (200 and 400GbE switches and cables are expected to arrive towards the end of 2018.)

Technical computing is also driving the network need for speed. This includes fields such as oil and gas exploration; genomic research and new drug discovery; computer chip design; weather forecasting and climate modeling; automotive and aerospace design; artificial intelligence training and inferencing; and the editing/production of high-resolution video. While cloud deployments range from thousands to tens of thousands of servers at a time, technical computing clusters range from as few as 8 up to several thousand servers. Many of these deployments have traditionally used non-Ethernet interconnects such as InfiniBand or Intel OmniPath Architecture for design, research and simulation workloads, or Fibre Channel and the Serial Digital Interface (SDI) for video editing, but as they require faster speeds, many of these workloads are moving to 25Gb, 40Gb, 50Gb and 100Gb Ethernet.

Enterprises and Storage Vendors Follow Soon After

Enterprises often lag behind the cloud in adopting new IT paradigms, so it's not surprising that these enterprises are moving more slowly to adopt 25GbE and 100GbE. The networking vendor sales teams are showing up with shiny new 25/100GbE adapters and switches, just as many enterprise are finishing their upgrades to 10GbE client connections and 40GbE uplinks. But while most enterprise networking today is still at 10GbE and 40GbE in the datacenter (1GbE to the desktop and 1/2.5/5GbE for wireless), it's clear that the largest enterprises are also starting to deploy 25GbE to the server and 100GbE uplinks to aggregation switches or to the core routers.

Enterprise storage vendors have also lagged but are rapidly moving towards faster speeds. Today about half of them support 40GbE and/or 32G Fibre Channel (FC), while the other half are still shipping 10GbE and 8/16G FC ports. By mid-2019 it's likely all of them will support 40GbE and 32G FC, while the more innovative storage vendors will also be shipping 25GbE, with 100GbE support for the faster all-flash arrays.



To simplify enterprise deployments, nearly all 25GbE adapters and switch ports are backwards compatible with 10GbE, while 50GbE and 100GbE adapters and switch ports are backwards compatible with 40GbE. (32/128G FC equipment is also backwards compatible with 8/16G FC.) Modern Ethernet switches can usually support a mix of 10Gb, 25Gb, 40Gb, and 100Gb speeds on the same switch. Customers are starting to roll out 25/100GbE equipment for all new network infrastructure knowing it can support legacy servers, storage, and uplinks with 10/40GbE speeds as well as new servers, storage and uplinks with 25/100GbE speeds.

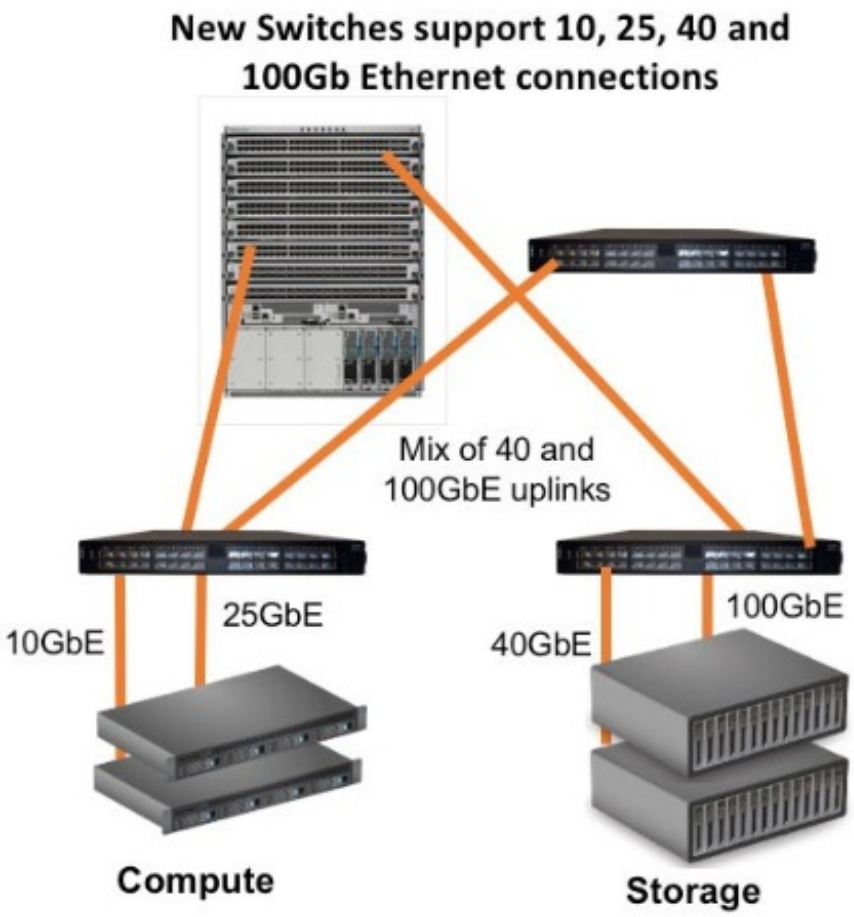


Figure 3: New Ethernet switches are backwards-compatible, supporting both 10/40GbE connections to older infrastructure as well as 25/100GbE connections to newer servers, storage and switches.

Where Were You When the Revolution Happened?

The 25/100GbE Revolution is well underway. It's already happening—and nearly done in some cases—amongst the cloud service providers, underway for technical computing, and in the late planning stages for enterprise and enterprise storage vendors. Due to the speed advantages, backwards compatibility, and near-cost-parity of 25G/100GbE over 10/40GbE, it makes financial sense for customers and vendors to choose 25GbE to the server and 100GbE for switch uplinks and some flash storage. Designing networks that can support these higher speeds helps create a datacenter infrastructure that is future-proof and ready for faster servers, faster storage, and new applications such as AI, Big Data, and hyperconverged infrastructure. Customers, vendors and system integrators should all consider the advantages of a 25/100GbE network infrastructure for new build-outs and network upgrades.

About the SNIA Ethernet Storage Forum

The SNIA Ethernet Storage Forum (ESF) is committed to providing vendor-neutral education on the advantages and adoption of Ethernet storage networking technologies. Learn more at http://www.snia.org/esf.