10GBASE-T for Broad 10 Gigabit Adoption in the Data Center

Responding to the increasing use of virtualization and unified networking by scaling Gigabit Ethernet increases network complexity and cost. Moving to 10 Gigabit Ethernet addresses these problems.

EXECUTIVE SUMMARY

Advancing server, storage, and unstructured data technologies are driving the need for data center bandwidth well beyond the Gigabit Ethernet (GbE) industry norm. Trends such as server consolidation, cloud virtualization, and Hadoop clustering present significant bandwidth challenges. Many customers address these challenges by doubling or quadrupling the number GbE interfaces per device. This approach, while preserving investments in cabling and operation best practices, adds complexity and limits scalability.

10 Gigabit Ethernet (10GbE) addresses these limitations while providing multiple cabling options, including 10GBASE-T, which uses familiar twisted-pair copper cabling and RJ-45 interfaces. The latest generation of 10GBASE-T interfaces reduces the average watt per Gigabit to less than that of GbE connections. And while the cost of a 10GbE port is in the range of two to three times that of a GbE port, as shown in Figure 1, the total cost of ownership is less when one factors in the cost, complexity, limited scalability, cabling, and rack space associated with doubling or quadrupling GbE connections.

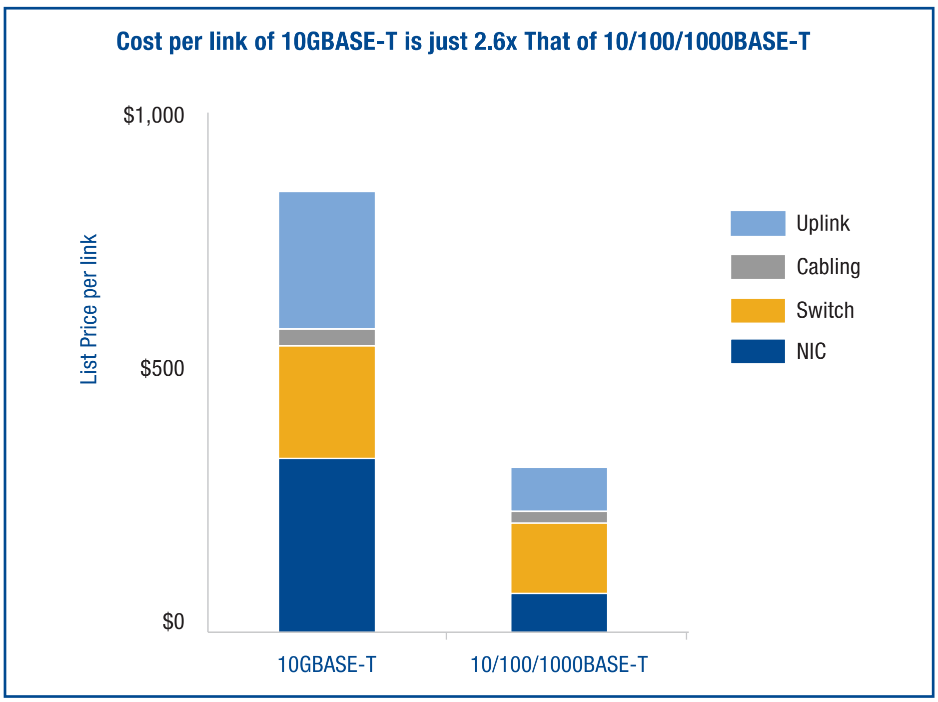

Figure 1. Cost comparison between 10GBASE-T and 1000BASE-T[1]

Solution provided by:

CONTENTS

Broad deployment on 10GBASE-T will simplify data center infrastructures, making it easier to manage server and storage connectivity while delivering the bandwidth needed for heavily virtualized servers and I/O-intensive applications.

10 Gigabit Ethernet: Drivers for Adoption

The growing use of virtualization in data centers to reduce IT costs has caused many administrators to take a serious look at 10GbE as a way to reduce the complexities they face when using existing GbE infrastructures. The server consolidation associated with virtualization has had significant impact on network I/O because it combines the network needs of several physical machines and background services, such as live migration, which can easily consume up to eight gigabits per second of bandwidth, onto the I/O of a single machine.

Further, storage trends, including unified LAN and SAN networking and Big Data solutions such as the Hadoop Distributed File System (HDFS), are increasing I/O demands to the point where a GbE network can be a bottleneck and a source of complexity in the data center. Implementing a common transport for all of these storage types requires a rethinking of the entire data center infrastructure, and while GbE connections might be able to handle the bandwidth requirements of a single traffic type, they do not have adequate bandwidth for multiple traffic types during peak periods. Network designers either need to add many more GbE interfaces or come up with a different yet compatible approach.

Moving to 10GbE addresses these network problems by providing greater bandwidth, reducing latency, freeing up precious rack space, and simplifying cabling. A single 10GbE interface delivers more the twice the bandwidth of four GbE interfaces, with significant consolidation benefits. Data Center Administrators have a number of 10GbE interfaces to choose from, including CX4, SFP+ Fiber, SFP+ Direct Attach Copper (DAC), and 10GBASE-T. Today, the majority of 10GbE networking deployments use SFP+ Optical or SFP+DAC interfaces. Many have chosen these interfaces based upon lower power consumption, lower latency, and optical interfaces' support for longer distances when compared to the copper options on the market.

Much has changed in 2012 specific to 10GBASE-T and the benefits of third-generation physical layer controller (PHY) silicon. This latest generation of 10GBASE-T PHYs significantly lowers the power requirements per port and offers backwards compatibility with GbE cabling infrastructures, per the IEEE 802.3an specification.

Widespread deployment of 10GbE requires a cost-effective solution that is backwards-compatible with existing networks and has the distance to reach the majority of switches and servers in the data center. This white paper looks at what is driving adoption of10GbE and how recent developments in 10GBASE-T, including LAN on motherboard (LOM) integration, will drive broader deployment of 10GbE. It also outlines the advantages of 10GBASE-T, including improved bandwidth, greater flexibility, infrastructure simplification, ease of migration, and cost reduction.

The Need for 10 Gigabit Ethernet

A variety of technological advancements and trends are driving the increasing need for 10GbE in the data center. For instance, increasingly powerful Intel® Xeon® processors are boosting server performance, which allows customers to host more applications on a single server. Customers are also using virtualization to consolidate multiple servers onto a single physical server, maximizing server utilization while reducing their equipment, floor space, and power costs.

However, server consolidation and virtualization have a significant impact on a server's network bandwidth requirements, as the I/O needs of what were several servers now need to be met by a single physical server's network resources. To match the increase in network I/O demand, many IT organizations have scaled their networks by doubling, tripling, or even quadrupling the number of GbE connections per server. This model has led to increased networking complexity and costs, as it requires additional Ethernet adapters, network cables, switch ports, switches, and rack space. Moreover, there are added layers of software complexity, as increasing the number of interfaces, while load sharing across these interfaces requires link aggregation control (LACP) between the server and switch interfaces. These complex configurations require more sophisticated network engineers.

The transition to unified networking adds to the increasing demand for high bandwidth networking. IT departments are moving to unified networking to help simplify network infrastructure by converging LAN and SAN traffic, including iSCSI and NAS. This convergence drives greater network consolidation and simplifies the network, but it also significantly increases network I/O demand by enabling multiple traffic types to share a common Ethernet network.

Continuing down the GbE path is not sustainable, as the added complexity, power demands, and cost of additional GbE adapters will not allow customers to scale to meet current and future I/O demands. Simply put, scaling GbE to meet these demands significantly increases the cost and complexity of the network. Moving to 10GbE addresses the increased bandwidth needs while greatly simplifying the network and lowering power consumption by replacing multiple GbE connections with a single or dual-port 10GbE connection.

Media Options for 10 Gigabit Ethernet

Despite strong growth in 10GbE deployments, 10GbE has yet to achieve mass market Data Center adoption, due to a number of limitations, primarily around cost, power consumption, and cabling options. The following section outlines several of these limitations in greater detail and discusses how the current generation of 10GBASE-T technologies removes them going forward.

10GBASE-CX4

10GBASE-CX4 was an early favorite for 10 GbE deployments, but its adoption was limited by bulky and expensive cables and limited reach of 15 meters. The large size of the CX4 connector, when compared to other interface options, prohibited higher switch port densities, specifically with 1RU top of rack switches and modular switch chassis. Moreover, pathways and spaces were not sufficient to handle the larger cables.

SFP+

SFP+'s support for both fiber optic cables and DAC make it a more flexible solution than CX4. SFP+ connectors are smaller than CX4 connectors and offer comparable port densities to their GbE counterparts. SFP+ DAC is the leading 10GbE interface today, but its limitations in reach and cost will prevent this media from becoming the de facto cabling and interface technology for 10GbE.

10GBASE-SR (SFP+ Fiber)

Cabling with Fiber is great for latency and distance (up to 300 meters). Fiber cabling coupled with optical transceivers offers the best power consumption footprint, however it is more expensive than other 10GbE media types; optical transceivers can add up to 30-40% to server, switch, and storage interface costs. The fiber electronics can be four to five times more expensive than their copper counterparts, meaning that ongoing active maintenance, typically based on original equipment purchase price, is also more expensive. This drives up both the acquisition costs as well as the ongoing annual maintenance contracts. Further, the vast majority of SFP+ connections are sold as add-in server adapters, as opposed to 10GBASE-T connections, which are now integrated on server motherboards. This add-in model adds cost and maintenance overhead.

10GBASE-SFP+ DAC

DAC is a lower cost alternative to fiber, but its reach for passive cables is limited to 7 meters, and it is not backwards-compatible with existing GbE switches. DAC requires the purchase of an add-in adapter and uses a new top of rack (ToR) switch topology. DAC cables are much more expensive than structured copper channels and cannot be field-terminated, making DAC more costly than 10GBASE-T. DAC cables carry a 90 day warranty, whereas a structured cabling system carries a 20 year warranty. The adoption rate of DAC for LAN on Motherboard (LOM) will be low, since it does not have the flexibility and reach of 10GBASE-T. With top of rack deployments, it is very difficult to use all the switch ports purchased due the generally lower number of server adapter ports and the limited reach of the cables. These unused ports carry an initial cost outlay and require power (even in idle mode) and maintenance costs, making them expensive on an ongoing basis.

10GBASE-T

10GBASE-T offers the most flexibility, the lowest cost media, and is backwards-compatible with existing GbE networks.

Reach

Like all BASE-T implementations, 10GBASE-T supports lengths up to 100 meters, giving IT managers a far greater level of flexibility in connecting devices in the data center and accommodating top of rack, middle of row, or end of row network topologies. 10GBASE-T also gives IT managers the most flexibility in server placement, since it works with existing structured cabling systems.

For higher grade cabling plants (category 6A and above), 10GBASE-T operates in low power mode (also known as data center mode) on channels under 30 m. This means a further power savings per port over the longer 100m mode. Data centers can create any-to-all patching zones to assure less than 30m channels to realize these savings.

Backwards Compatibility

Because 10GBASE-T is backward-compatible with 1000BASE-T, it can be deployed in existing GbE switch infrastructures that are cabled with Category 6 and Category 6A (Cat 6 and Cat 6A) or greater cabling, helping IT to keep costs down while offering an easy migration path to 10GbE. Unlike SFP+ options, a 10GBASE-T connection can auto-negotiate and auto-select the proper port speed when plugged into a GbE port.

Power

Early 10GBASE-T PHYs consumed too much power for widespread adoption. The same was true when GbE products were released, with first-generation GbE chips using roughly 6.5 Watts per port. With process improvements, the chips improved from one generation to the next, and today's GbE ports use less than 1 W per port. The same has proven true for 10GBASE-T; PHYs are Moore's Law-friendly, and newer process technologies have reduced both the power and cost of the latest generation of 10GBASE-T PHYs.

When 10GBASE-T adapters were introduced in 2008, they required 25 W of power for a single port. Power has been reduced in successive generations, using newer and smaller process technologies. Intel's latest 10GBASE-T controller, the Intel® Ethernet Controller X540, requires less than 6.5 Watts per port, making it suitable for LOM integration. Today's third-generation PHYs have enabled switch ports to show similar improvements, with the latest Arista 7050T switches using just 6.5 watts per switch port. And while this is two to three times the power use of GbE ports, 10GBASE-T offers triple the benefit in power consumption, and 5X benefit in the cost per port (when looking at aggregate bandwidth consumption factors).

Latency

Depending on packet size, latency for 1000BASE-T ranges from below 1 µs to over 12 µs. 10GBASE-T ranges from just over two µs to less than 4 µs – a much tighter latency range. Latency for the Arista 7050T Switch, for example, is 3.3 µs and remains relatively flat across all packet sizes. For Ethernet packet sizes of 512 bytes or larger, 10GBASE-T's overall throughput offers an advantage over 1000BASE-T. Latency for 10GBASE-T is more than three times lower than 1000BASE-T with larger packet sizes. When one considers that software and application latency contributes much higher latency overall than that of 10GBASE-T, it is clear that only the most latency-sensitive applications such as High Performance Computing (HPC) or high frequency trading systems would be affected by normal 10GASE-T latency. These applications are a small fraction of the overall applications that broad deployment of 10GBASE-T technology will address.

The incremental 2 µs latency of 10GBASE-T when compared to 10GbE SFP+ is of no consequence to most users. For the large majority of enterprise applications that have been operating for years with 1000BASE-T latency, 10GBASE-T latency only makes things better. Many LAN products intentionally add small amounts of latency to reduce power consumption or CPU overhead. For example, consider interrupt moderation, a common LAN feature. Enabled by default, this feature typically adds ~100 µs of latency in order to allow interrupts to be coalesced and greatly reduce the CPU burden. For many users, this trade-off provides an overall positive benefit.

Cost

As power consumption has dropped dramatically over the last three generations, cost has followed a similar downward curve. First-generation 10GBASE-T adapters cost $US 1000 per port. Today's fourth-generation, dual-port 10GBASE-T adapters cost less than $US 300 per port. Integration as LOM in servers based on the Intel Xeon processor E5 product family has driven costs down further, integrating cost in the overall price of the server. By using 10GBASE-T LOM connections, users will see a significant savings over the purchase price of more expensive SFP+ DAC and fiber optic adapters, which cost several hundred dollars per adapter, and will be able to free up an I/O slot in the server. 10GBASE-T switch prices have also dropped considerably, with per-port prices at less than $350.

Data Center Network Architecture Benefits of 10GBASE-T

Table 1 lists the typical data center network architectures supported by various switching platforms and form factors. The table shows that switches with 10GBASE-T ports offer greater design flexibility when compared to

Table 1. Data Center Deployment Models for 10 GbE

TECHNOLOGY	DATA CENTER NETWORK ARCHITECTURE	CONNECTIVITY
10GBASE-SR(SFP+ Fiber)	• Top of Rack (ToR)	Uplinks from ToR switches to aggregation layer switches
	• Middle of Row (MoR)	Inter-cabinet connectivity from servers to MoR switches
	• End of Row (EoR)	Inter-cabinet connectivity from servers to EoR switches
	• Core network	Backbone
	• IDA/HDA/MDA	Intermediate, Horizontal and Main distribution areas as outlined in TIA 942-A
	• MD/ZD/EO	Main Distributor, zone distributor and Equipment Outlets as outlined in ISO 24764
10GBASE-SFP+ DAC	Top of Rack	Intra-cabinet connectivity from servers to ToR switches
10GBASE-CX4	Top of Rack	Intra-cabinet connectivity from servers to ToR switches
10GBASE-T	• Top of Rack (ToR)	Uplinks from ToR switches to aggregation layer switches Intra-cabinet connectivity from servers to ToR switches
	• Middle of Row (MoR)	Intra-cabinet connectivity from servers to MoR switches
	• End of Row (EoR)	Inter-cabinet connectivity from servers to EoR switches
	• IDA/HDA/MDA	Intermediate, Horizontal and Main distribution areas as outlined in TIA 942-A
	• MD/ZD/EO	Main Distributor, zone distributor and Equipment Outlets as outlined in ISO 24764

SFP+ DAC active and passive copper cabling alternatives. This flexibility is based on greater distance capabilities, lower cabling and port costs, and simpler moves, adds, and changes. Further, 10GBASE-T switches support a greater number of deployment options, including top of rack, end of row, middle of row, and inter-cabinet connectivity.

Switch port densities also need to be factored in when designing and specifying a data center network architecture. Customers typically purchase switches based on both the density of the server and storage racks and the bandwidth (number and speed of ports) required to interconnect these racks at wire speed. Smaller rack configurations require switches that offer between 24 to 36 ports, while larger rack configurations require switches that range up to 48.

In the future as server processing and Ethernet-based storage continue to grow, the number of 10 Gbps port connections required will expand beyond 48 ports within a rack. Again, 10GBASE-T is well-positioned here due to lower cabling costs, the advancements of lower powered PHYs, and the broader distance options for deployment.

The Future of 10GBASE-T

Many in the industry foresee broad deployment of 10 GbE in the form of 10GBASE-T. As shown in the Figure 2 below, in 2010 SFP+ (Fiber and Direct Attach Copper inclusive) represented 34% of the 10 GbE physical media in data centers. This percentage is projected to continue to rise through 2013 to 58%, with large deployments in IP Data Centers and for High Performance Computing. 10GBASE-T will grow from only one percent of physical media in 2010 to 17% in 2013 and eventually become the predominant media choice with 61% of 10GbE deployments in 2016.

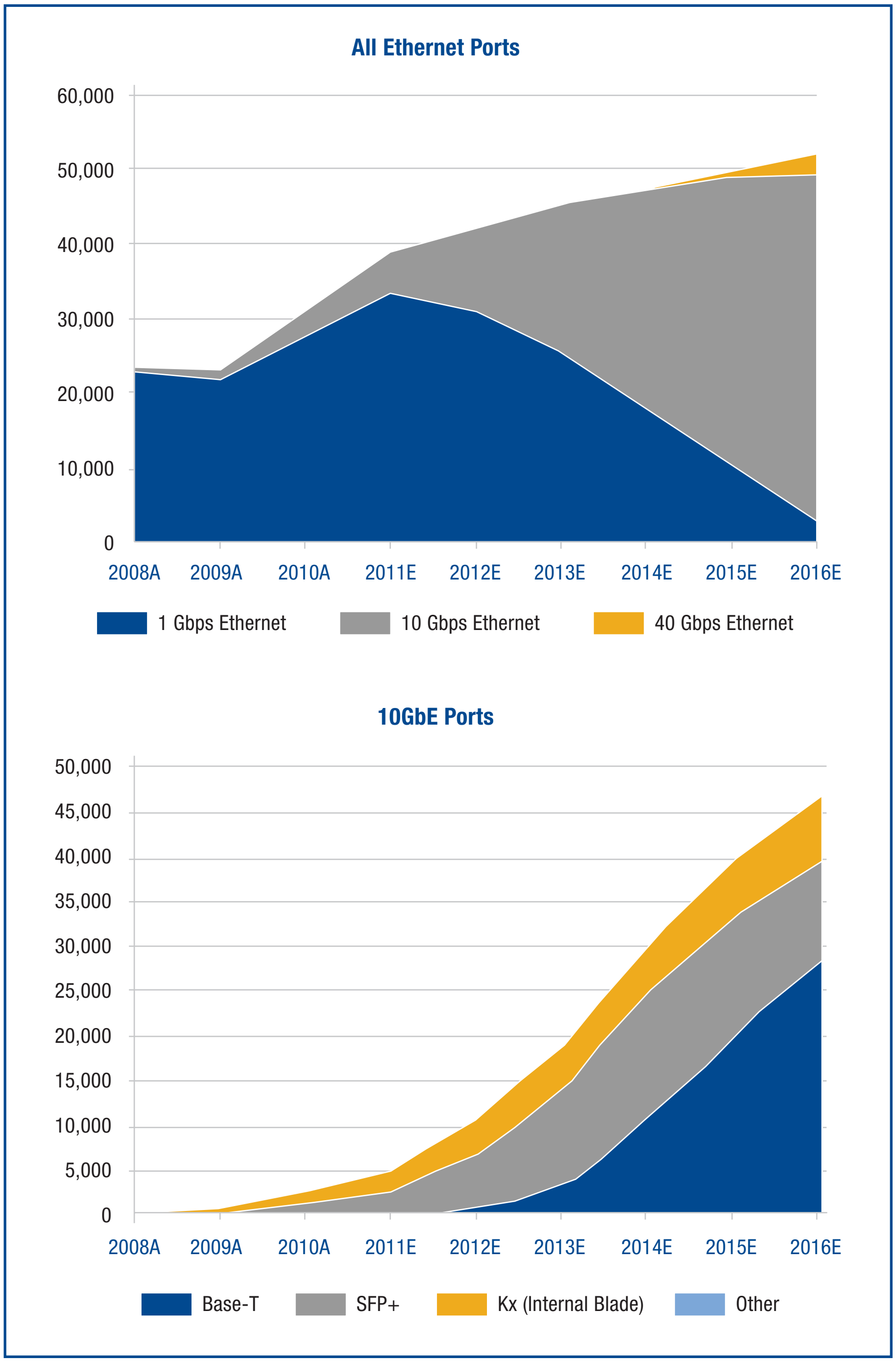

Figure 2. Projected Ethernet growth and 10GbE interface mix[2]

10GBASE-T as LOM

Integrated 10GBASE-T LAN on Motherboard (LOM) connections are featured on many servers based on the new Intel® Xeon® processor E5 family. With 10GBASE-T, OEMs can create a single motherboard design to support GbE, 10 GbE, and any distance up to 100 meters. 1000BASE-T is the incumbent in the vast majority of data centers today, and 10GBASE-T is the natural next step.

Conclusion

Broad deployment on 10GBASE-T will simplify data center infrastructures, making it easier to manage server and storage connectivity while delivering the bandwidth needed for heavily virtualized servers and I/O-intensive applications. Third generation advancements in 10GBASE-T PHYs have significantly lowered the cost per port, as well as the power consumption for 10GbE. These advancements, new 10GBASE-T switches such as the Arista 7050T series, and single-chip MAC/PHY designs, such as the Intel® Ethernet Controller X540, are now driving 10GBASE-T integration on the server motherboard. 10GBASE-T switches that offer significant port density (up to 64 ports per rack unit) and significantly reduced power per port requirements are readily available.

Clearly a perfect storm has emerged for the broad deployment of 10GBASE-T networks: strong growth in bandwidth-hungry applications has driven the need for increased server I/O, powerful new server processors are driving data centers consolidation and virtualization, deployment of Ethernet-based storage systems is increasing, and a new generation of 10GBASE-T PHY technology has lowered costs and power consumption to make 10GBASE-T LOM a reality. These factors and key 10GBASE-T features, including backwards-compatibility with GbE infrastructures and distance flexibility, will help drive 10GBASE-T to a prominent place in the data center.

For more information, visit:

www.intel.com/go/ethernet

www.aristanetworks.com

www.commscope.com

www.siemon.com

---

[1] Source: CommScope

[2] Crehan Research, Server-class Adapter and LOM, January 2012

 XXXX/XXX/XXX/XX/XX XXXXXX-001US