ethernet alliance

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

Contributors

Carl G. Hansen, Intel
Carrie Higbie, Siemon
Yinglin (Frank) Yang, Commscope, Inc

November 2010

Table of Contents

10 Gigabit Ethernet: Drivers for Adoption

The growing use of virtualization in data centers to address the need to reduce IT costs has caused many administrators to take a serious look at 10Gb Ethernet (10GbE) as a way to reduce the complexities they face when using the existing 1Gb Ethernet (1GbE) infrastructures. The server consolidation associated with virtualization has had significant impact on network I/O because they combine the network needs of several physical machines and the other background services, such as live migration, over the Ethernet network onto a single machine.

Together with trends such as unified networking, the ability to use a single Ethernet network for both data and storage traffic, are increasing I/O demands to the point where a 1GbE network can be a bottleneck and a source of complexity in the data center. The move to implement unified networking requires rethinking of data center networks. While 1GbE 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. This creates a need for multiple 1GbE connections.

Moving to 10 Gigabit Ethernet (10GbE) addresses these network problems by providing more bandwidth and simplifies the network infrastructure by consolidating multiple gigabit ports into a single 10 gigabit connection. 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, most are choosing either 10GbE Optical or SFP+ DAC. However, limitations with each of these interfaces have kept them from being broadly deployed across the data center.

Fiber connections are not cost-effective for broad deployment, and SFP+ DAC is limited by its seven meter reach and requires a complete infrastructure upgrade. CX4 is an older technology that does not meet high density requirements. For 10GBASE-T, the perception to date has been that it required too much power and was too costly for broad deployments. These concerns are being addressed with the latest manufacturing processes that are significantly reducing both the power and cost of 10GBASE-T.

Widespread deployment requires a cost-effective solution that is backward compatible and has the flexibility capable of reaching the majority of switches and servers in the data center. This white paper looks at what is driving choices for deploying 10GbE and how 10GBASE-T will lead to broader deployment, including its integration into server motherboards. It also outlines the advantages of 10GBASE-T in the data center, 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, the widespread availability of multi-core processors and multi-socket platforms is boosting server performance. That performance allows customers to host more applications on a single server resulting in multiple applications competing for a finite number of I/O resources on the server. Customers are also using virtualization to consolidate multiple servers onto a

single physical server reducing their equipment and power costs. Servers using the latest Intel® Xeon® processors can support server consolidation ratios of up to fifteen to one[1].

However, server consolidation and virtualization have a significant impact on a server's network bandwidth requirements as the I/O needs of several servers now need to be met by a single physical server's network resources. To match the increase in network I/O demand, IT has scaled their network by doubling, tripling, or even quadrupling the number of gigabit Ethernet connections per server. This model has led to increased networking complexity as it requires additional Ethernet adapters, network cables and switch ports.
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, NAS, and FCoE for a single Ethernet data center protocol. This convergence does simplify the network but significantly increases network I/O demand by enabling multiple traffic types to share a single Ethernet fabric.

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 gigabit connections with a single or dual port 10GbE connection.

[1] Source: Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

Media Options for 10 Gigabit Ethernet

Despite industry consensus regarding the move to 10GbE, the broad deployment of 10GbE has been limited due to a number of factors. Understanding this dynamic requires an examination at the pros and cons of current 10GbE media options.

Technology	Pros	Cons	Cable
10GBASE-SR (SFP+ Fiber)	• 3.2w / port • Latency: < 1 µs • 300 meter reach (SR) • 10km reach (LR) • < $699 / port	• Requires new SFP+ switches • High cost of purchase and deployment • Not compatible with installed 1000BASE-T switches • Due to the higher purchase price ongoing maintenance costs are higher	• Multimode fiber
10GBASE-DAC (SFP+ Direct-Attach Copper)	• 2.9w / port • Latency: < 1 µs • < $399 / port	• Limited to 7 meters • Not compatible with existing 1000BASE-T switches • Requires new SFP+ switches • Can lead to a significant oversubscription of ports due to limited reach	• Twinax copper
10GBASE-CX4	• 2.25w / port	• Limited to 15 meters • Large connector form factor limits high density • Requires 10GBASE-CX4 switches, with few new products available • Not compatible with existing 1000BASE-T • Larger diameter cabling is difficult to manage in legacy cabinets	• Twinax copper
10GBASE-T	• <$399 / port • 9.2w / port (<80m) • 10.1w / port (>80m) • CAT6 55m • CAT6a 100m • CAT7 100m • Backward compatibility with 1000BASE-T • Uses existing high performance structured cabling	• Higher power than SFP+ • Latency: 2 - 4 µs	• UTP copper • FTP copper

Table 1: 10GbE Media Options

The challenge IT managers face with 10GbE currently is that each of the current options has a downside, whether in terms of cost, power consumption, or reach.

10GBASE-CX4

10GBASE-CX4 was an early favorite for 10GbE deployments; however, its adoption was limited by the bulky and expensive cables and its reach is limited to 15 meters. The size of the CX4 connector prohibited higher switch densities required for large scale deployment. Larger diameter cables are purchased in fixed lengths resulting in challenges to manage cable slack. Pathways and spaces may not be sufficient to handle the larger cables.

SFP+

SFP+'s support for both fiber optic cables and DAC make it a better (more flexible) solution than CX4. SFP+ is ramping today, but has limitations that will prevent this media from moving to every server.

10GBASE-SR (SFP+ Fiber)

Fiber is great for latency and distance (up to 300 meters) but it is expensive. Fiber offers low power consumption, but the cost of laying fiber networking everywhere in the data center is prohibitive due largely to the cost of the electronics. The fiber electronics can be 4-5 times more expensive than their copper counterparts meaning that ongoing active maintenance, typically based on original equipment purchase price, is also more expensive. Where a copper connection is readily available in a server, moving to fiber creates the need to purchase not only the fiber switch port, but also a fiber NIC for the server.

10GBASE-SFP+ DAC

DAC is a lower cost alternative to fiber but it can only reach 7 meters and it is not backward compatible with existing GbE switches. DAC requires the purchase of an adapter card and requires a new top of rack (ToR) switch topology. The cables are much more expensive than structured copper channels and cannot be field terminated. This makes DAC a more expensive alternative to 10GBASE-T. The adoption of DAC for LOM will be low since it does not have the flexibility and reach of BASE-T.

10GBASE-T

10GBASE-T offers the most flexibility, is the lowest cost media type, and is backward compatible with existing 1GbE networks.

Reach

Like all BASE-T implementations, 10GBASE-T works for lengths up to 100 meters giving IT managers a far-greater level of flexibility in connecting devices in the data center. With flexibility in reach, 10GBASE-T can accommodate either top of the rack, middle of row, or end of the row network topologies. This gives IT managers the most flexibility in server placement since it will work 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 30m. This equates to 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 this savings.

Backward Compatibility

Because 10GBASE-T is backward-compatible with 1000BASE-T, it can be deployed in existing 1GbE switch infrastructures in data centers that are cabled with CAT6, CAT6A or above cabling, allowing IT to keep costs down while offering an easy migration path to 10GbE.

Power

The challenge with 10GBASE-T is that the early physical layer interface chips (PHYs) have consumed too much power for widespread adoption. The same was true when gigabit Ethernet products were released. The original gigabit chips were roughly 6.5 Watts/port. With process improvements, chips improved from one generation to the next. The resulting GbE ports are now under 1W / port. The same has proven true for 10GBASE-T. The good news with 10GBASE-T is that these PHYs benefit greatly from the latest manufacturing processes. PHYs are Moore's Law friendly and the newer process technologies will continue to reduce both the power and cost of the latest 10GBASE-T PHYs.

When 10GBASE-T adapters were first introduced in 2008, they required 25w of power for a single port. Power has been reduced in successive generations of using newer and smaller process technologies. The latest 10GBASE-T adapters require only 10w per port. Further improvements will reduce power even more. By 2011, power will drop below 5 watts per port making 10GBASE-T suitable for motherboard integration and high density switches.

Latency

Depending on packet size, latency for 1000BASE-T ranges from sub-microsecond to over 12 microseconds. 10GBASE-T ranges from just over 2 microseconds to less than 4 microseconds, a much narrower latency range. For Ethernet packet sizes of 512B or larger, 10GBASE-T's overall throughout offers an advantage over 1000BASE-T. Latency for 10GBASE-T is more than 3 times lower than 1000BASE-T at larger packet sizes. Only the most latent sensitive applications such as HPC or high frequency trading systems would notice any latency.

The incremental 2 microsecond latency of 10GBASE-T 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 gets better. Many LAN products purposely add small amounts of latency to reduce power consumption or CPU overhead. A common LAN feature is interrupt moderation. Enabled by default, this feature typically adds ~100 microseconds 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 metrics have dropped dramatically over the last three generations, cost has followed a similar downward curve. First-generation 10GBASE-T adapters cost $1000 per port. Today's third-generation dual-port 10GBASE-T adapters are less than $400 per port. In 2011, 10GBASE-T will be designed as LAN on Motherboard (LOM) and will be included in the price of the server. By utilizing the new resident 10GBASE-T LOM modules, users will see a significant savings over the purchase price of more expensive SFP+ DAC and fiber optic adapters and will be able to free up and I/O slot in the server.

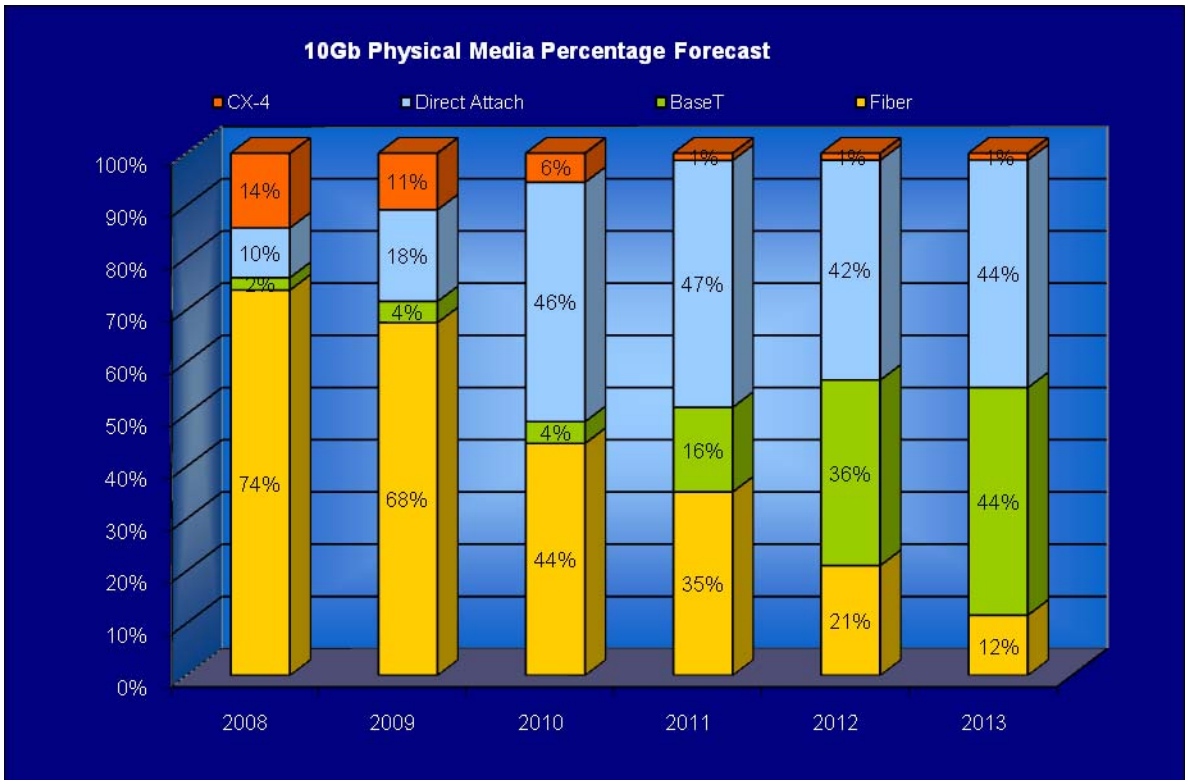
Data Center Network Architecture Options for 10 Gigabit Ethernet

The chart below lists the typical data center network architectures applicable to the various 10GbE technologies. The table clearly shows 10GBASE-T technology provides greater design flexibility than its two copper counterparts.

Technology	Data Center Network Architectures	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
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)	Intra-cabinet connectivity from servers to ToR 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

The Future of 10GBASE-T

Intel sees broad deployment of 10GbE in the form of 10GBASE-T. In 2010, fiber represents 44% of the 10GbE physical media in data centers but this percentage with continue to drop to approximately 12% by 2013. Direct-attach connections will grow over the next few years to 44% by 2013 with large deployments in IP Data Centers and for High Performance Computing. 10GBASE-T will grow from only 4% of physical media in 2010 to 44% in 2013 and eventually becoming the predominate media choice.



Source: Intel Market Forecast

10GBASE-T as LOM

Sever OEMs will standardize on BASE-T as the media of choice for broadly deploying 10GbE for rack and tower servers. 10GBASE-T provides the most flexibility in performance and reach. OEMs can create a single motherboard design to support GbE, 10GbE, and any distance up to 100 meters. 1GBASE-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 connectivity while delivering the bandwidth needed for heavily virtualized servers and I/O-intensive applications. As volumes rise, prices will continue to fall and new silicon processes have lowered power and thermal values. These advances make 10GBASE-T suitable for integration on server motherboards. This level of integration, known as LAN on Motherboard (LOM) will lead to mainstream adoption of 10GbE for all server types in the data center.

About Ethernet Alliance

The Ethernet Alliance is a community of Ethernet end users, system and component vendors, industry experts and university and government professionals who are committed to the continued success and expansion of Ethernet. The Ethernet Alliance brings Ethernet standards to life by supporting activities that span from incubation of new Ethernet technologies to interoperability demonstrations, certification and education.