

Upgrading the Data Center to 10 Gigabit Ethernet

The standard for 10 Gigabit Ethernet (IEEE802.3ae) was ratified in 2002. While 10GbE deployments have grown every year since then, the technology has primarily been used to interconnect switches and routers. Almost all of the server connections in data centers have remained at 1 Gbps, limiting the amount of network throughput available to each server. With recent enhancements in CPU performance, system I/O, and storage I/O the gigabit network has increasingly become the application and workload performance bottleneck.

The primary reason for staying with Gigabit Ethernet has been cost-performance. Until recently it has been more cost-effective to have multiple GbE connections rather than a single 10 GbE port. In addition, most installed servers typically cannot utilize the full bandwidth of a 10 GbE connection. However both of these factors are changing, which will lead to widespread adoption of 10 GbE for server connectivity over the next few years.

Increased Demand for Network Performance

The demand for network performance is driven by increasing application demands, specifically:

- Next generation multi-core CPUs with multi-threaded networking stacks will be able to fully utilize a 10 Gbps connection. Server CPU throughput has historically doubled every 18 months, and there is every reason to expect this trend to continue.
- The increase in CPU throughput drives increased demands in storage I/O. The fastest growing storage architecture are so-called unified storage servers that offer both network file storage as well as network block storage, a natural fit for 10 Gigabit Ethernet. The shift from rotating media to solid-state storage architectures is also delivering a 100x increase in storage IOPS that can benefit from a higher performance network.
- Server virtualization is increasing server utilization and therefore the demands for network bandwidth per physical server. Gone are the days of large numbers of servers running at low utilization rates, and with this increased server utilization, comes increased network bandwidth needs. Technologies like VM mobility and VM fault tolerance also keep all workload state synchronized between disparate servers requiring constant high throughput and low-latency networking.
- Massively parallel cluster computing requires high bandwidth and low-latency between all the nodes in the cluster. 10 Gigabit Ethernet offers a 10X improvement in this area compared to 1 Gigabit Ethernet.
- Real-time applications such as distribution of financial market data demand the lowest latency with scalable multicast. 10 Gigabit Ethernet is the natural solution for this requirement.

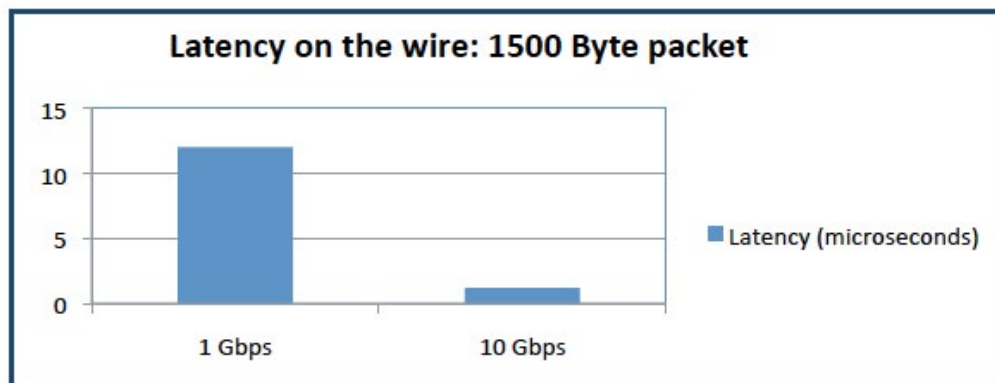
Scalability of the Network Fabric

A 10 Gigabit data center network offers many advantages in terms of scalability, including the following:

- Enables terabits of aggregate traffic, without adding more layers to the network
- Simplifies the network design by eliminating congestion points
- Reduces the need for complex QoS schemes
- Supports server virtualization that increases network demands due to server consolidation
- Improves application efficiency, server utilization, and power efficiency
- Avoids the need to use multiple 1 GbE NICs and connections to get higher bandwidth

Lower Latency

10 Gigabit Ethernet offers a compelling improvement in end-to-end latency. For HPC applications and financial trading environments, low latency is critical. Not only does 10 GbE reduce the serialization delay by a factor of 10x, but many 10 GbE switches also deliver an order of magnitude lower latency than 1 GbE switches. For latency sensitive application, 10 Gigabit Ethernet is a major improvement.



Reliability

Many data centers require 24x7 network availability and have no provisions for scheduled downtime. To support this on the network side requires a higher level of hardware and software reliability design than the conventional enterprise design practice. Fortunately, next generation 10 Gigabit Ethernet switches offer fundamental reliability improvements for both fixed/rack (1RU) and modular switches, including:

- Hot-swappable redundant power and cooling
- Highly modular software design to offer fault containment
- Stateful Fault Containment (SFC) and Stateful Fault Repair (SFR)
- In service software upgrades (ISSU) for the entire network

Density, Power and Cooling of 10GbE Switches

Density, Power and Cooling are key drivers for deployment of data center network infrastructure.

- The latest 10 GbE top of rack switches now support the same 48-port density as 1 GbE switches, which means that the customer does not lose valuable rack space when upgrading to 10 GbE.
- Next generation 10 GbE top of rack switches use front-to-rear or rear-to-front cooling to match the server airflow. This is essential for efficient cooling of switches located in server racks and retaining the flexibility to use them as part of a modular switch/router port expansion or structured cabling plant design maintaining ASHRAE compliance.
- Power consumption per 10 GbE port is still a significant concern, however the latest 10 GbE Switches have much improved power consumption compared to earlier models with an overall lower power draw when compared to the throughput.
- Power consumption for 10GBASE-T is higher than 10 GbE ports utilizing transceivers. This will continue to decrease as the chipsets continue to evolve over the next several generations of 10GBASE-T.

Connectivity

How to connect servers over 10 GbE has continued to evolve.

- 10GBASE-T switches began to enter the market in late 2009. This advancement allows for the connectivity of devices over 10 GbE using traditional Unshielded Twisted Pair cabling. Given that most servers today remain connected over Gigabit Ethernet running over Category 5/6/6A Unshielded Twisted Pair cabling, 10GBASE-T provides a convenient and highly cost effective solution to migrate servers to 10 GbE without necessarily requiring the installation of new cabling or upgrading the servers at the same time.
- A large number of 10 GbE connections from servers to a Top of Rack switch, switch to switch interconnects, or storage arrays to a 10 GbE switch are frequently over fairly short distances (.5m – 7m is common place). With twinax copper cabling (CX-1), these interconnects can be made very cost effectively for devices not yet 10GBASE-T capable.
- 10 GbE optical transceivers remain commonplace and 2010 will bring new optical offerings, which will deliver longer reach solutions (> 10Km) in addition to the existing options available today.

Challenges in moving to 10GbE

While 10 GbE offers many advantages, there are certain challenges as well.

- Not all switches offer mixed-speed ports. As a result, you are forced to upgrade all your servers and the network at the same time. In many cases, this is hard to do operationally and/or financially.
- 10 GbE NICs are still somewhat expensive. This will change once 10 GbE NICs are integrated on future server mainboards (Frequently referred to as LAN on Mainboard (LOM), which is expected to start shipping during 2010.

- 10 GbE optical transceivers are still expensive, and only some server and networking vendors offer twinax copper cables or 10GBASE-T as more cost effective solutions.

The cost per 10 GbE port, including the server NIC and the physical layer, remains the greatest impediment to large-scale adoption of 10 GbE for servers. During 2010, the cost for a 10 GbE server connection is expected to decline below the cost of multiple GbE server connections, which is the present approach to scale server bandwidth beyond 1 Gigabit Ethernet.

Future Trends

It is clear that the migration to 10 Gigabit Ethernet has started. New data center build outs everywhere are considering the advantages of 10 GbE. The percentage of servers directly attached to 10 GbE has gone from approximately 8% in 2008 to roughly double that in 2009. Analysts predict 10 GbE attached servers to double again in 2010 and by 2012 it is expected that the majority of all new servers will be connected directly via 10 GbE.

The IEEE 802.3 high-speed task force is now defining the standard for 40 and 100 Gigabit Ethernet, which is a key development for future backbone connectivity. It is expected that the first switches that support 40 and 100 Gigabit Ethernet will become available in 2011.

Summary

For next-generation data centers, 10 Gigabit and higher speed Ethernet is the interconnect of choice, providing key improvements in terms of bandwidth, latency, scalability, reliability and application performance. If you are planning to upgrade your data center fabric, you should consider the latest 10 GbE solutions that are cost-effective compared to multiple Gigabit Ethernet ports but deliver substantially better throughput and latency.

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



Copyright © 2016 Arista Networks, Inc. All rights reserved. CloudVision, and EOS are registered trademarks and Arista Networks is a trademark of Arista Networks, Inc. All other company names are trademarks of their respective holders. Information in this document is subject to change without notice. Certain features may not yet be available. Arista Networks, Inc. assumes no responsibility for any errors that may appear in this document.