Addressing the Needs and Concerns of Higher Speeds of Ethernet

A Tehuti Networks white paper

How to contact us

For inquiries regarding Tehuti Networks, please use the information below.

Email:

info@tehutinetworks.net

Website:

www.tehutinetworks.net

Version 1.6

January 2007

Executive Summary

As the industry moves towards 10 Gigabit Ethernet (10GbE) there is an increasing demand for greater performance and capabilities. Although the market has seen the introduction of 10GbE products into the market for a few years, these products have not been able to provide full 10Gbps performance and have been quite expensive. However, new standards and architectures are being introduced to the market that will allow the industry to not only achieve much better performance from 10GbE solutions, but to do so at much lower price points.

The Need for 10 Gigabit Ethernet

Today, the data center is preparing for a volume transition from 1 Gigabit Ethernet (GbE) to 10 Gigabit Ethernet (10GbE). This transition results from the fact that simultaneously the market is demanding the adoption of higher speed Ethernet from several different directions:

- Increasing network traffic, both in terms of the number and size of transactions due in part to the Internet and increased personal connectivity
- The growth of increasing data-intensive applications, such as grid and cluster computing
- Service offerings such as On Demand Viewing, IPTV, broadband and video streaming
- Increasing bandwidth to residential and business clients
- The limitations of existing protocols, such as link aggregation (LAG), to handle the growth.

One of the key drivers for 10GbE is the limitation of using the link aggregation protocol to aggregate multiple 1GbE links. LAG is frequently used as an interim step to bridge the gap between the existing speed of Ethernet technology and the next higher speed of Ethernet technology. However, it is only a temporary solution as more than four physical links becomes very difficult to manage and costly to implement and maintain. For example, a standard network interface card (NIC) only has enough room on its faceplate for four connections. Also, the LAG protocol has overhead associated with each additional link; therefore, with every link that is added to the LAG group, the overhead will increase slightly on each link already in the group. 10GbE has the advantage of providing higher bandwidth via a single link and offers lower latency due to the elimination of the LAG protocol and its overhead. In addition, 10GbE offers a wide range of cable runs – from 15 meters on coax cables to 40 kilometers on single-mode fiber (SMF).

Although the market has enjoyed the existence of 10GbE products such as switches and adapters for approximately five years - the standard for 10GbE over fiber-optic transmission media (IEEE Std. 802.3ae™) was ratified in 2002 - there have been impediments that have prevented 10GbE from being widely adopted. With the ever-increasing need for higher bandwidth, these obstacles will be overcome within the next few years to permit 10GbE to penetrate into volume platforms.

The higher speed and faster response time of 10GbE can benefit any traffic-burdened network. However, 10GbE is particularly beneficial to virtually any data-intensive application. Examples include enterprise financial applications, database and data modeling simulations, weather forecasting, computer-aided design and manufacturing, and graphics intensive applications such as those often found in computer games and films. Newer applications such as RFID is beginning to drive huge databases of information across networks and cellular networks are moving to broadband. GbE found its way in to high-performance computing (HPC) cluster applications and as the costs and barriers to entry decline, 10GbE will grow in this market. Some of the more controversial applications include the possibility that at 10 Gb/s speeds, other technologies such as iSCSI begin to make economic as well as technical sense, potentially replacing existing protocols such as fibre channel.

Metro area network (MAN) applications as well as long distance local area network (LAN) applications also greatly benefit from the longer distances available via fiber optics. An example is the ability to place datacenters where it is most advantageous in terms of cost and convenience, such as offsite from the primary facilities. The longer reach of 10GbE and its ability to map to existing SONET/SDH 10Gb/s protocols also makes it very suitable for wide area network (WAN) and Internet points of presence (POP) applications to implement a network capable of transmitting data at terabyte-per-hour rates.

From an end user perspective there are many common consumer activities that drive the need for additional bandwidth. Digital photos from 7-10 megapixel cameras are now frequently housed using online storage. High resolution video downloads of 2-3 GB each, automatic software upgrades occur online and video games such as Xbox™ and PlayStation®3 are offered with live music and video feeds. All of these types of activities result in the need for bigger and faster servers that must be serviced by fatter network pipes.

What 10GbE is and is Not

In many ways 10GbE is a straightforward extension of previous Ethernet speeds allowing it to be fully backward compatible to previous Ethernet generations. It retains the key Ethernet architecture; the media access controller (MAC) protocol, the Ethernet frame format, and the minimum and maximum frame size. Figure 1, below, provides an overview of the relationship of 10GbE relative to the OSI layer model.

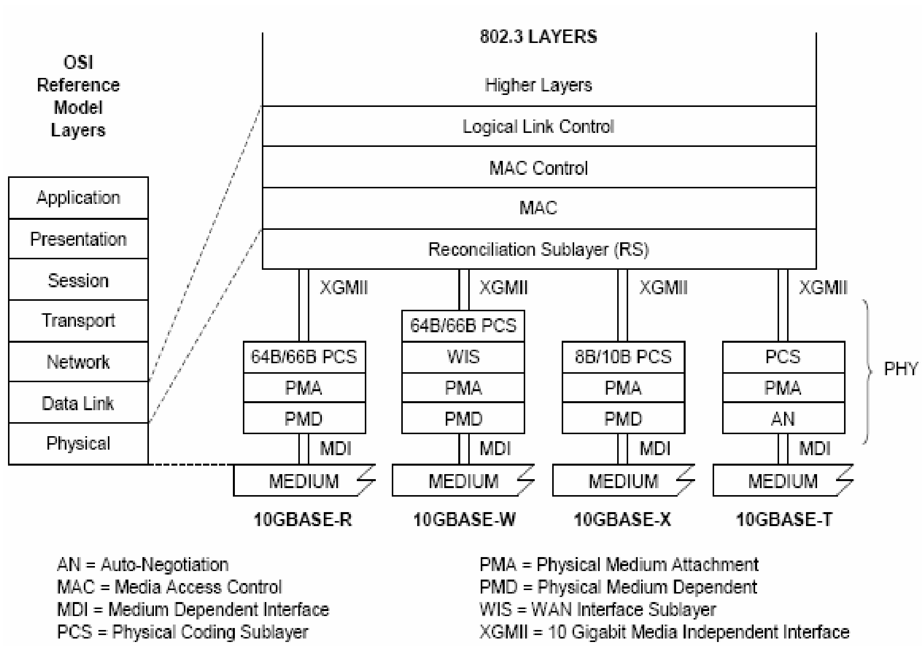

Figure 1 – 10GbE relationship to OSI layer model

However, there are also some key differences that primarily involve the physical layer (PHY) of the Open Systems Interconnection (OSI) layer model (Figure 1). While the same MAC and RS are used for all 10GbE devices, the PHYs used to connect to the network transmission media are different. In addition, unlike previous generations of Ethernet speeds, 10GbE operates only in full-duplex mode. By 1999, the use of switched Ethernet had grown in the market to the point that support for half-duplex mode of operation was considered a burden to the development of 10GbE.

The original 10GbE standard (IEEE Std. 802.3ae™-2002) specified a variety of PHYs for connection to multimode fiber (MMF) and SMF transmission media. Since the original standard ratification, several new amendments to the 10GbE standard have been developed or have started development. In 2004, IEEE Std. 802.3ak™-2004 specified a 10GBASE-CX4 PHY for the transmission of 10GbE signaling over 15 meters of coax cabling for short reach applications. In 2006, two new PHYs were added (10GBASE-T and 10GBASE-LRM), and another two will be added in 2007 (10GBASE-KR and 10GBASE-KX4). The new PHYs address the expanding market needs for 10GbE.

The table (Table 1) below provides an overview of the standards and draft standards for 10GbE's various transmission media.

Standard or Draft	Ratification Date	PHY types	Medium	Reach (max)
Std. 802.3ae	2002	10GBASE-SR	MMF	300m
		10GBASE-LR	SMF	10km
		10GBASE-ER	SMF	40km
		10GBASE-LX4	MMF/SMF	300m/10km
		10GBASE-SW	MMF	300m
		10GBASE-LW	SMF	10km
		10GBASE-EW	SMF	40km
Std. 802.3ak	2004	10GBASE-CX4	Coax	15M
Std. 802.3an	June 2006	10GBASE-T	Cat 6 UTP or better	100M
Std. 802.3aq	Sept 2006	10GBASE-LRM	MMF	220M
P802.3ap	March 2007*	10GBASE-KX4	4-lane FR backplane	1M
		10GBASE-KR	1-lane FR4 backplane	1M

***** Expected ratification date

Table 1 – 10GbE Standards and Draft Standards

Even though two 10GbE standards have existed for a number of years, the cost of the 10GbE optical solutions and the short reach of 10GbE coax solution have limited the widespread adoption of 10GbE. The cost and power of optics solutions continue to decrease as the technology matures, but the lack of a standard that allowed 10GbE to operate over unshielded twisted-pair (UTP) copper cabling relegates 10GbE to applications where it is critical. The June 2006 ratification of the 10GBASE-T standard (IEEE Std. 802.3an™- 2006) addressed the lack of a UTP solution and the market will start to see lower cost 10GbE copper PHYs appear.

The 10GBASE-T standard has some common misperceptions associated with it. There is much confusion around the support of auto-negotiation, but the 10GBASE-T standard does indeed support auto-negotiation. The auto-negotiation feature will permit implementers to design PHY devices that operate over a broader range of speeds and media. While auto-negotiation will provide the capability to be backwards compatible to lower speeds, it is unlikely that a 10GBASE-T system will be able to support operation at 10 Mb/s due to the requirements of the transformer.

Another common misperception is around the support for Category 5 (CAT-5) cabling. CAT-5 and CAT-5e cabling are not listed as supported media in IEEE Std. 802.an™ - 2006, due to the fact that the performance requirements for 10GBASE-T far exceed the specified performance of the cabling. The 10GBASE-T standard calls for the support of CAT-6, CAT-6 Augmented (CAT-6A) and CAT-7 cabling. Existing CAT-6 installations meeting the requirements of Telecommunications Industry Association (TIA – www.tiaonline.org) telecommunications systems bulletin 155 (TSB-155) should support a reach of up to 55m. CAT-6A and CAT-7 support a reach of up to 100m. The primary difference between CAT-6A and CAT-7 is that CAT-6A is an unshielded cable, whereas, CAT-7 is shielded.

From a board consideration, the optimal interconnect will be primarily determined by the reach required. For example, 10GBASE-CX4 provides a low-cost, higher performance solution for short-haul applications. It is being used as a high performance interconnect for clients involved in data intensive applications and as a replacement for proprietary fabrics in storage area networks (SANs).

For networks that require longer-haul distances such as MANs and WANs, 10GbE offers the ability to reach up to 40 km. This provides large potential benefits in terms of providing much faster and more efficient access for applications such as telecommuting, video conferencing, data mining and online research. In WAN applications, 10GbE allows the transport of terabytes of data in world-record breaking times, providing the global linking of communities. The ability of 10GbE to effectively attach directly to the SONET/SDH core network or to be transported over unused wavelengths (aka dark lambda) minimizes or eliminates the need for protocol translation for global connectivity.

What is Needed to Achieve 10GbE Speeds?

Implementing 10GbE to achieve true 10 Gb/s speeds requires more than simply choosing the right medium. Much of the 10GbE system throughput has to occur both in the infrastructure (such as routers and switches) as well as the endpoints (primarily servers, appliances and networking storage equipment).

To meet these demands, several technical and economic concerns must be met. Increasingly more expensive and higher performance systems are required to keep pace with network throughput yet due to design constraints, the memory subsystem cannot physically keep up. This means that latency is increasing as processors wait through hundreds of idle cycles while memory works at its slower rate to move data across the network to the waiting applications. The memory subsystem and processors are literally being flooded by the network bandwidth, so the industry has been seeking ways to alleviate this bottleneck. This disparity is impacting the performance of today's systems designed to handle critical real-time transactions and bandwidth intensive applications like e-Commerce, medical imaging and data warehousing.

One of the most optimal approaches to solving this bottleneck problem is using a host hardware offload solution. A controller that offloads the TCP/IP protocol, such as that offered by Tehuti Networks (Figure 2 and Figure 3), reduces the flooding by optimizing the ability of the system to process the packets. Implemented from wire to application, accelerator technology dramatically speeds up enterprise platform to network communications by reducing latency, redistributing functions to components that will do each task better, and using additional enhancements as appropriate to improve server system throughput. This allows the servers to maintain an optimal balance of performance and low-power while requiring minimal board real estate; thereby, reducing the total cost of ownership (TCO) of the equipment.

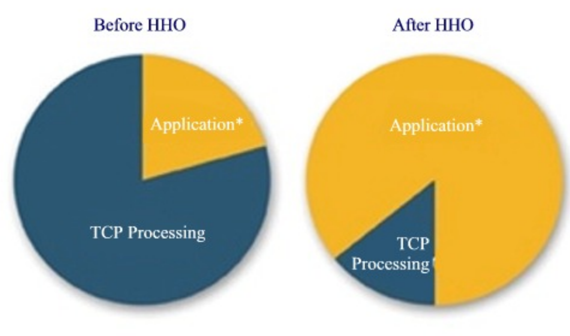

Figure 2 – Freeing up CPU cycles with a Host Hardware Offload (HHO)

In addition, interfaces are an important consideration. For example, the 64-bit PCI-X bus pumps out data in the multi-gigabit range, whereas the PCI Express™ architecture easily scales to handle 10GbE.

Figure 3 – Tehuti Networks 10 GbE host hardware offload controllers, shown here mounted on a network interface card, reduce latency by optimizing the system's ability to process packets

What Comes After 10GbE?

Although the industry is in the throes of preparing for the ramp of 10GbE into volume, history tells the industry it is already time to begin thinking about how a higher speed Ethernet will come to fruition. The aforementioned concerns will only be further exacerbated and new concerns added as the market looks towards the next speed generation after 10GbE. It is predicted by some that by 2012 the next higher speed Ethernet adoption could mirror that of 10GbE and approach approximately 50% share of the primary interconnect in the Top 500 supercomputing sites. In July 2006, it is anticipated there will be a Call for Interest (CFI) in IEEE 802.3 for a Higher Speed Study Group (HSSG), which could result in the ratification of the next higher speed Ethernet standard sometime in the 2010 timeframe.

Summary

10GbE shares many similarities with previous generations of Ethernet, but also has some key differences. In the 10GbE standard itself, many of the differences are related to the transmission media, but in order to achieve true 10GbE performance, the system must be configured to support 10GbE throughout the network. This includes everything from improving TCP/IP processing via hardware assist so the CPU has more cycles for application processing, to making sure the right interfaces (e.g. PCI Express) are in place, to having chipsets with memory subsystems that won't limit the system performance. Many of these issues are being addressed today, and the market can look forward to enjoying true 10GbE performance and all its benefits in the very near future.

About Tehuti Networks

Tehuti Networks is a fabless semiconductor company offers high-performance solutions-on-a-chip (SoC) for offloading TCP/IP in applications that need to achieve near 10Gb/s performance, but are sensitive to the power, real estate, and cost. The solutions improve server processing performance and provide significant cost benefits to original equipment manufacturers and IT users. More information is available at **www.tehutinetworks.net**.