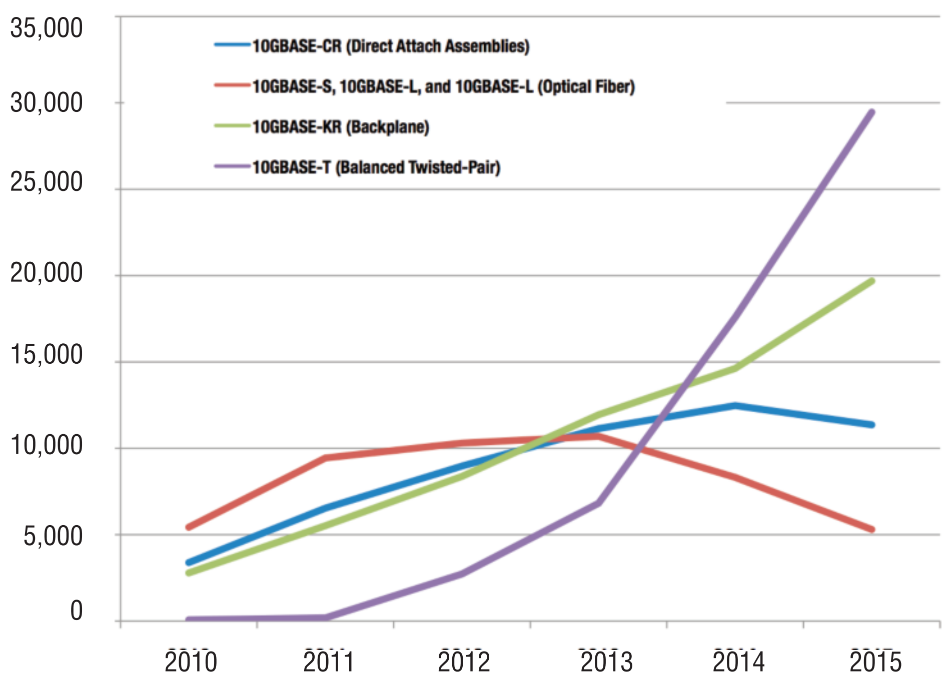
State of the Network: 10GBASE-T Equipment Availability and the Future of Copper Media

While it may be so that good things come to those who wait, too much waiting can lead to uncertainty. Take 10GBASE-T networking products, for example. The 10GBASE T Standard[1] published almost six years ago and the long wait for network gear has provided fodder for the digital rumor mill to churn. This had led to the completely erroneous misperception that 10GBASE-T is the end of the line for copper balanced twisted-pair media and network equipment. The fact is that the extended time to market can be explained by the recent economic recession and the desire to integrate significant power efficiency enhancements into this new technology. These challenges have been overcome and all indicators are that adoption of 10GBASE T solutions is poised to take off in 2012.

This paper presents the truths behind the myths surrounding 10GBASE-T and the future of copper twisted-pair Ethernet applications.

10GBASE-T network equipment is available and deployment rates are increasing. Although initially hampered by power-hungry implementations, today's chip technology that delivers the 10GBASE-T bit stream (also called a "PHY") capitalizes on an advanced 40nm lithography manufacturing process, which cuts power use, board size, and cost. As a result, significant adoption of 10GBASE-T technology is expected to begin in 2012. During this year, at least 20 new platforms (e.g., switches, servers and NICs) using 10GBASE-T PHY devices are expected to have broad market availability. In addition, a new market research report issued by The Linley Group[2] forecasts over 2.7 million ports of 10GBASE-T PHYs to ship in 2012 -- a sharp rise from the 182,000 ports counted as shipped in 2011. The trend lines shown in Figure 1 depict The Linley Group's forecast for several different types of 10 Gb/s Ethernet applications over the next few years. Note that 10GBASE-T is expected to achieve a dominant market position in 2014. The adoption rates forecasted in Figure 1 are consistent with the historical Ethernet adoption profile whereby optical networking interfaces initially precede copper interfaces but copper port counts greatly outnumber optical port counts soon thereafter.

Figure 1: Forecasted adoption rate of 10 Gb/s Ethernet Applications



Source: The Linley Group

10GBASE-T and copper balanced twisted-pair cabling offer unique benefits compared to other 10 Gb/s Ethernet solutions. With cost and power dissipation significantly reduced with the newer 40nm PHY devices, and further reductions enabled by 28nm devices expected in 2013, data center managers can now capitalize on the fundamental advantages offered by 10GBASE-T technology that include:

- the ability to interoperate with legacy slower-speed Ethernet technologies though the function of auto negotiation
- the ease of deploying a copper balanced twisted-pair cabling system and the use of familiar cabling and connector interfaces
- the flexibility of 100-meter, 4-connector structured cabling topologies to support additions, moves, and changes in LAN and data center environments, and
- the ability to deliver Power over Ethernet (PoE and PoE Plus)

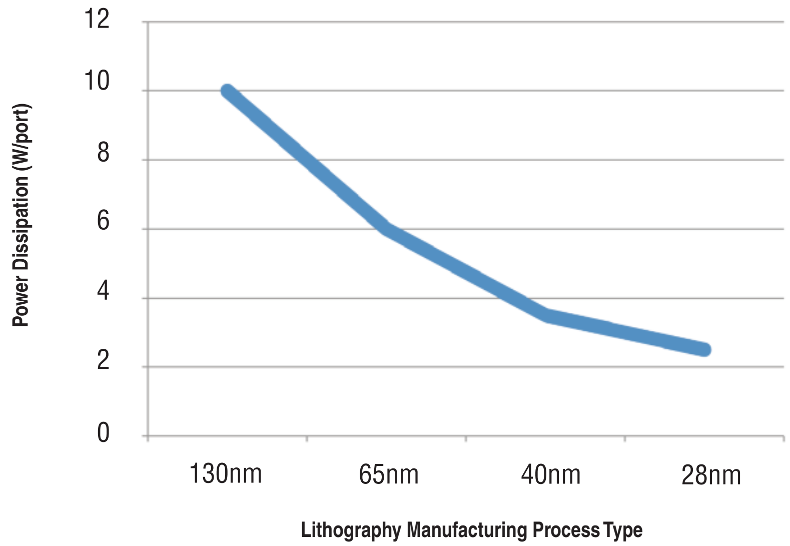
This IC device from PLX Technology represents the latest generation of 10GBASE-T transceivers. Implemented in 40nm lithography, it can be purchased in single channel, dual channel, and quad channel versions.

Interoperability with legacy Ethernet equipment via auto-negotiation is of particular significance as it enables data center expansions and expenditures to occur incrementally. Rather than demanding a wholesale upgrade of all servers and switches to 10 Gb/s capability, which is necessary for non-negotiating Ethernet systems transmitting over optical fiber media or direct attach assemblies such as SFP+, 10GBASE T network equipment supports 10 Gb/s transmission to new servers and can also auto-negotiate down to 1 Gb/s (or slower) speeds to support legacy servers. In this way, data centers can “future proof” their switching architectures. A 10GBASE-T switch can communicate effectively with legacy 1 Gb/s and 100 Mb/s servers today and allow 10 Gb/s servers to be introduced when required and supported by expense allocations tomorrow.

Another major catalyst to 10GBASE-T adoption will be the introduction of LAN-on-Motherboard (LOM) chips. Expected to be introduced in mid-2012, these devices will allow server manufacturers to also implement auto negotiation technology into their gear. The implications of this development are quite profound as for the first time servers will come preconfigured with Ethernet connections able to negotiate to 100 Mb/s, 1 Gb/s, or 10 Gb/s speeds depending on the capabilities of other devices in the network. The data center manager will want to be ready for this development by deploying 10GBASE-T capable switches that can extract the full capability of the server that it is connected to.

10GBASE-T PHY power consumption is well managed. 10GBASE-T device power dissipation has been closely scrutinized and has been rapidly declining since the technology was first standardized by the IEEE 802.3 in 2006. Early PHY implementations were created using a 130nm lithography manufacturing process and dissipated approximately 10W per port. By comparison, the 40nm devices available today are capable of less than 4W per port dissipation. And, the 28nm devices anticipated in 2013 are expected to dissipate less than 2.5W per port! Figure 2 highlights this trend.

Figure 2: 10GBASE-T Power Dissipation



Two special protocols can further improve 10GBASE-T power dissipation. In addition to the reductions afforded by advances in semiconductor technology, BASE-T systems, and 10GBASE-T systems in particular, are able to take advantage of some unique and standards-based algorithms that exploit the nature of computer traffic to further reduce power dissipation. In particular, there are opportunities to improve efficiency when network equipment is idle for both sustained and very short time periods.

Wake-on-LAN (WoL) is a new networking standard formed by the Advanced Manageability Alliance whereby network equipment, such as a server, is put in sleep mode until awakened by a special network signal called a “magic packet.” The server’s network interface card (NIC) reverts to a very low power-dissipation mode during the sleep period, but remains alert and waiting for the magic packet. Once the packet arrives, the server is awakened and normal operation is resumed. Since the wakeup time associated with WoL is typically tens of seconds, this power management strategy is best suited for use when servers are expected to be idle for long periods of time -- such as at night or during other lengthy periods of inactivity. Even the most active of data centers experiences periods of time in which only a portion of its capacity is needed. This is a natural consequence of over-building resources to accommodate peak compute demands and the temporal and seasonal fluctuation in those demands due to non-uniform user locations and time schedules. WoL can take advantage of these demand fluctuations with startling results; putting even a single server with a typical power dissipation of 500W to sleep gains much more benefit than the difference in power of hundreds of transceiver devices! Equally important, 10 Gb/s Ethernet deployed over optical media or SFP+ direct attach assemblies is not designed to support the WoL protocol at this time and, as a result, these systems always dissipate their full power. WoL is an important strategy employed uniquely by 10GBASE-T to reduce overall power consumption in the data center.

While WoL is designed for lengthy idle periods, another technology called Energy Efficient Ethernet[3] (EEE) is specifically designed to take advantage of the bursty nature of computer traffic. It is the case that typical Ethernet traffic contains many gaps, which can range in duration from microseconds to milliseconds. Heretofore, these gaps have been filled with so-called "idle patterns" or waveforms containing no real computer information, but whose transitions can be used for maintaining clock synchronization between transceivers. The EEE algorithm exchanges those idle patterns for a Low Power Idle (LPI) mode where very little power is dissipated.

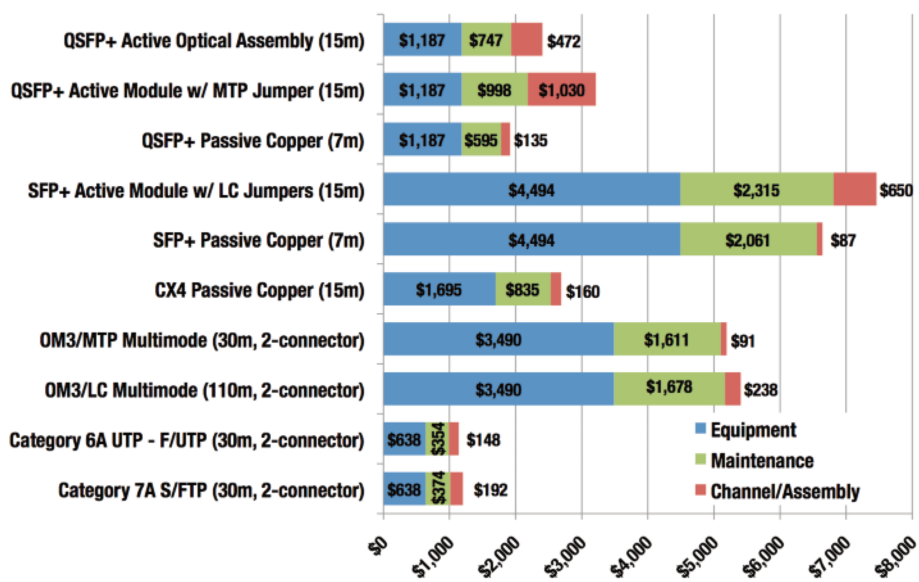
The LPI mode used during idle periods requires a new signaling scheme composed of alerts over the line and alerts to and from station management. When in the LPI mode, a refresh signal is used to keep receiver parameters, such as timing lock, equalizer coefficients, and canceller coefficients, current. These are also critical to enable fast transitions from LPI to Active modes. Typical transition times from Active to LPI mode and back are in the three-microsecond range, so implementation of EEE introduces minimal latency into the network. The bottom line is that transceiver power savings utilizing the EEE algorithm can range from 50 percent to 90 percent depending on actual data patterns. For example, a 28nm 10GBASE-T transceiver with a typical Active mode power dissipation of 1.5W for 30-meter reach will dissipate only 750mW when utilizing the EEE algorithm with typical computer data patterns. Even better, system-level EEE optimizations implemented in switches and Ethernet controller silicon are expected to save far more power than EEE optimizations in the transceiver because the energy consumption of the entire switch or server (which is more than double the power per port of even the previous generation of transceivers) can be leveraged.

10GBASE-T "Short Reach" mode is another 10GBASE-T power dissipation improvement strategy. Another feature present in 10GBASE-T PHYs, which can greatly aid in the reduction of overall power dissipation, is the ability to automatically detect channel length between compliant transceivers. When the channel length is less than 100 meters, 10GBASE-T transceivers are able to reduce their power dissipation while still maintaining fully compliant Bit Error Rate (BER) performance. This so-called "Short Reach" mode takes advantage of the larger signal to noise ratios present due to lower signal attenuation in short channels and the power dissipation reductions can be dramatic. For example, since the signal strength at the receiver is significantly larger if is it attenuated by only 10 meters of cabling as opposed to 100 meters of cabling, transmit power can be significantly reduced without adversely affecting BER. It is a common misperception that Short Reach mode is an on-off condition that is directly tied to a specific link length (e.g., 30 meters). In fact, the Short Reach mode power dissipation profile is contiguous and scalable versus length.

In Short Reach mode, not only can transmit power be reduced, but the number of filter taps used for echo cancellation and line equalization can also be curtailed and powered down internally in the device. As an example, a transceiver typically with 3.5W of power dissipation when connected to a 100 meter channel can exhibit power dissipation of only 2.5W when connected to a 30-meter channel or less than 2W when connected to a 10-meter channel. Since many newer data center configurations rely on shorter cable lengths than the maximum length of 100 meters, exploiting this feature is becoming growingly important.

10GBASE-T is the most cost effective 10 Gb/s Ethernet application. While reach, power consumption, and backwards compatibility are important considerations when selecting media, most designers will assert that cost significantly influences the decision making process. The truth is that 10GBASE-T offers more benefits and flexibility than other 10 Gb/s applications at the most favorable price point! Figure 3 shows the equipment (server port and NIC), media, and yearly maintenance costs for one channel and its corresponding 10 Gb/s port connections, which are representative of the types and lengths of media commonly deployed in data centers. The most economical choice for 10 Gb/s transmission is 10GBASE T network equipment in conjunction with category 6A UTP, category 6A F/UTP, or category 7A S/FTP balanced copper twisted-pair cabling. The same conclusion is reached when this analysis is repeated for channels and their corresponding port connections that represent the types and lengths of media commonly deployed in horizontal LAN cabling. It is this cost advantage that will drive the rapid adoption of 10GBASE-T in 2012.

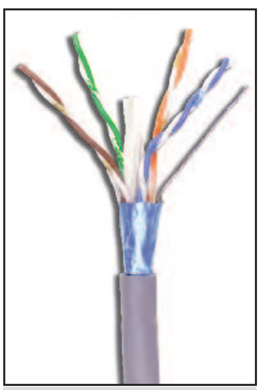
Figure 3: 10GBASE-T Cabling Cost Comparison



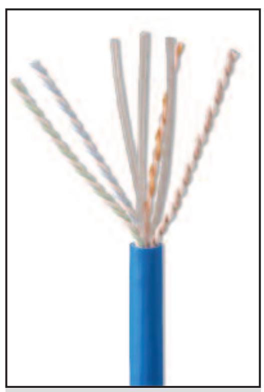
Interest in speeds beyond 10 Gb/s over copper balanced twisted-pair cabling is growing. The most significant confirmation that BASE-T Ethernet applications have a strong future is the growing interest in "Next Generation" cabling. This media will be targeted to support the copper balanced twisted-pair application that comes after 10GBASE-T! Because Ethernet applications for the LAN backbone and data center core have always preceded Ethernet specifications for the LAN horizontal and data center edge, it is a good bet that the next Ethernet over balanced twisted-pair speed will be 40 Gb/s to supplement IEEE 802.3ba[4]-compliant 40 Gb/s Ethernet computer backplanes and optical fiber network gear. At this time, the biggest driver demonstrating the great industry commitment to, interest in and investment in the future of copper-based Ethernet is the work that is being done by ISO/IEC and TIA to develop Next Generation cabling specifications to support such an application.

ISO/IEC recently initiated a project to develop a new standard tentatively titled "ISO/IEC 11801-99-x Guidance for balanced cabling in support of at least 40 GBit/s data transmission." This proposed two-part standard will address capabilities of both existing ISO/IEC 11801-compliant channels and channels with extended and/or enhanced performance characteristics. TIA is currently working on a project called "Specifications for 100Ω Next Generation Cabling," expected to be published as addendum 1 to ANSI/TIA-568-C.2. These massive project initiatives reaffirm the strength and popularity of BASE-T applications and balanced copper twisted-pair cabling media.

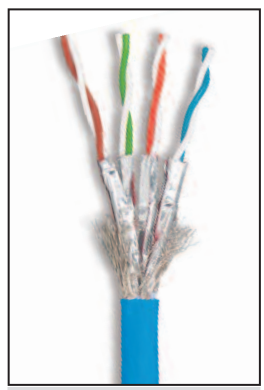
Figure 4: 10Gb/s Twisted-Pair Cable Options



Category 6A/Class E_A Shielded (F/UTP)



Category 6A/Class E_A UTP



Category 7A/Class F_A Fully-Shielded (S/FTP)

10GBASE-T was worth the wait. While 10 Gb/s Ethernet-ready copper balanced twisted-pair cabling has been available for some time, it has been a long and anxious wait for 10GBASE-T equipment to reach the broad market. That wait is over! 10GBASE-T network equipment offers greater reach and flexibility than any other 10 Gb/s copper solution and is a very attractive alternative to 10 Gb/s optical fiber solutions when deployed channel lengths are less than 100 meters. Data center and LAN IT managers who had the foresight to install 10 Gb/s Ethernet-ready copper balanced twisted-pair cabling in their network are poised to capitalize on the negotiation and power reduction features of 10GBASE-T and begin incremental server and switch upgrades to relieve network congestion and increase capacity this year. The rest have a little catching up to do.

About the Authors:

Ron Cates is vice president of marketing, networking products, at PLX Technology, a leader in 10GBASE-T transceivers. Prior to PLX, he was senior vice president and general manager of Wide Area Networking Products at Mindspeed Technologies. He has over 30 years of experience in the semiconductor industry and holds BSEE and MSEE degrees from the University of California at Los Angeles and an MBA from San Diego State University. He can be reached at rcates@plxtech.com.

Valerie Maguire, BSEE, holds the position of global sales engineer at Siemon, a world-class manufacturer of high performance copper and optical fiber cabling systems. She holds various leadership positions with the TIA TR-42 Telecommunications Cabling Systems Engineering Committee and IEEE 802.3 Ethernet Working Group. In addition, She has authored over 45 technical articles and engineering papers, holds one U.S. patent, and received the 2008 Harry J. Pfister Award for Excellence in Telecommunications. She can be reached at valerie_maguire@siemon.com.

References:

[1] IEEE 802.3at™, "IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment 1: Physical Layer and Management Parameters for 10 Gb/s Operation, Type 10GBASE-T", September 2006

[2] The Linley Group, "A Guide Ethernet Switch and PHY Chips", December 2011

[3] IEEE 802.3az™, "IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment 5: Media Access Control Parameters, Physical Layers, and Management Parameters for Energy Efficient Ethernet", October 2010

[4] IEEE 802.3ba™, "IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment 4: Media Access Control Parameters, Physical Layers, and Management Parameters for 40 Gb/s and 100 Gb/s Operation", June 2010

WP_State_of_10GBASE-T_ Rev. A 2/12

North America
Watertown, CT USA
Phone (1) 860 945 4200 US
Phone (1) 888 425 6165 Canada

Europe/Middle East/Africa
Surrey, England
Phone (44) 0 1932 571771

Asia/Pacific
Shanghai, P.R. China
Phone (86) 21 5385 0303

Latin America
Bogota, Colombia
Phone (571) 657 1950