ethernet alliance

10GBASE-T: 10 Gigabit Ethernet over Twisted-pair Copper

Version 1.0 August 2007

Authors[1]:
Hugh Barrass, Cisco
Michael Bennett, Lawrence Berkeley National Laboratory
Brad Booth, Applied Micro Circuits Corporation
Chris DiMinico, MC Communications
Paul Kish, Belden
David Law, 3Com
Shimon Muller, Sun Microsystems
Bruce Tolley, Solarflare Communications
George Zimmerman, Solarflare Communications

1. This work represents the opinions of the authors and does not necessarily represent the views of their affiliated organizations or companies.

Contents

List of Figures

List of Tables

Executive Summary

This white paper addresses three key topics about 10GBASE-T:

- The reason for 10GBASE-T, including a description of some 10GBASE-T applications;
- How it works, which explores the fundamental technology; and
- How you use it, which explains cabling and field-testing issues.

10GBASE-T promises to be a very successful high-speed networking solution for horizontal copper applications and high-performance networking in the following areas:

- HPC data centers both in new and existing installations;
- Enterprise server farms/data centers also with both new and legacy media
- Local uplinks, aggregation links and inter-switch links; and
- Other applications that can use in-building structured cabling with both new and legacy media.

In addition to describing the technology and applications of 10GBASE-T, this paper describes the link segment characteristics for all of the media supporting 10GBASE-T. It also explains how to test an existing installation and what can be done to improve cabling in preparation for 10GBASE-T.

Background

Copper cabling remains the medium of choice of network managers for horizontal cabling. More than 85 percent of the copper cabling currently inside buildings is Category 5e or better.[1]

Due to the continued predominance of copper cabling, top industry experts spent four years developing the IEEE Std 802.3an™-2006 standard so that the result offers the best cost performance trade-offs over the widest range of media. The

1. Flatman, A., LAN Technologies, IEEE Study Group, http://www.ieee802.org/3/10GBT/public/jan03/flatman_1_0103.pdf, 2003.

IEEE P802.3an task force defined a specification for running 10 Gigabit Ethernet over twisted-pair copper designated 10GBASE-T. By delivering 10GBASE-T networking solutions, the IT industry allows network managers to scale their networks to 10 Gigabit speeds while leveraging their investment in installed copper cabling infrastructure. In addition, for new installations, network managers can continue to leverage the cost-effectiveness and plug-and-play simplicity of copper structured cabling through the benefits of 10GBASE-T technology.

The strength of 10GBASE-T comes from its underlying technology. A section of this paper describes the line signaling; the coding; and the noise cancellation techniques necessary for operation on twisted-pair copper. The 10GBASE-T physical layer interface devices (PHYs) use sophisticated startup and power backoff mechanisms to ease the task of operating in the target environments. When it is implemented properly, these techniques become invisible to the user, but they are key to allowing 10GBASE-T to dramatically expand the applications for high-speed Ethernet.

The Reason for 10GBASE-T – Some Applications

This section describes some of the applications of 10GBASE-T, which illustrate both the reason and the need for its development. Optical 10 Gigabit and 10GBASE-CX4 which uses non-standard, short-reach, 15 meter InfiniBand™ cabling, are deployed today primarily for switch-to-switch links and switch-to-server links in high-performance computing clusters. This market, while growing, is small. Although 10G Ethernet has been shipping since 2001, in 2006 a relatively modest 300 thousand 10G optical and CX4 switch ports shipped. Compare this to the 5 to 6 million 1G copper switch ports (1000BASE-T) which ship each month.[2] 10GBASE-T switches and 10GBASE-T servers can dramatically expand the opportunity for 10G Ethernet networking by supporting simpler, cheaper, twisted-pair copper cabling necessary for the expansion of 10G networking beyond high-performance computing into commercial data centers and enterprise networks.

2. Dell'Oro Group, Ethernet Switching Report.

The Value of 10GBASE-T in the Data Center

Initial customer interest in 10GBASE-T is expected to come from network and IT managers with high-performance computing (HPC) requirements in data centers. The data center market can be segmented in many ways and the HPC networking environment includes Fibre Channel and InfiniBand equipment in addition to Ethernet. 10GBASE-T will accelerate deployment of 10 Gigabit Ethernet in data centers because it will cost less than optical 10 Gigabit Ethernet and be comparable in cost to a PHY/PMD InfiniBand copper port.

Plug-and-play Interconnect

There are many types of clusters used in data centers, some requiring engineered interconnects and others that take advantage of low-cost high-speed Ethernet. As shown in Figure 1, Gigabit Ethernet provides the majority of interconnects in the top 500 supercomputers.[3]

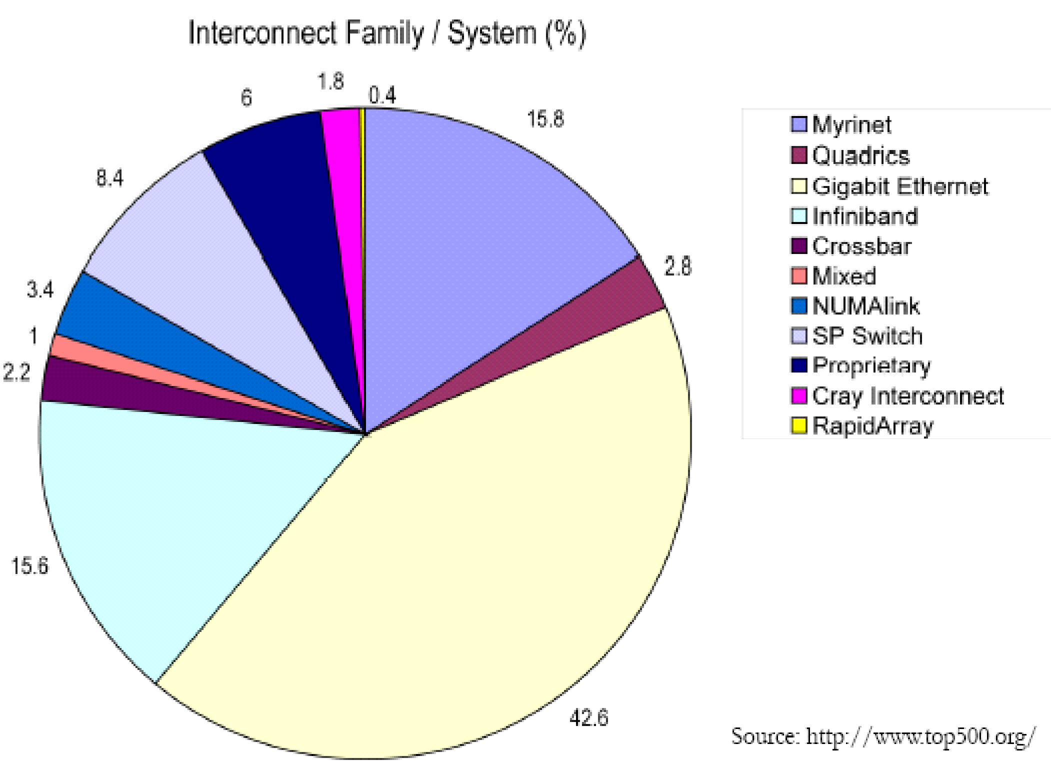


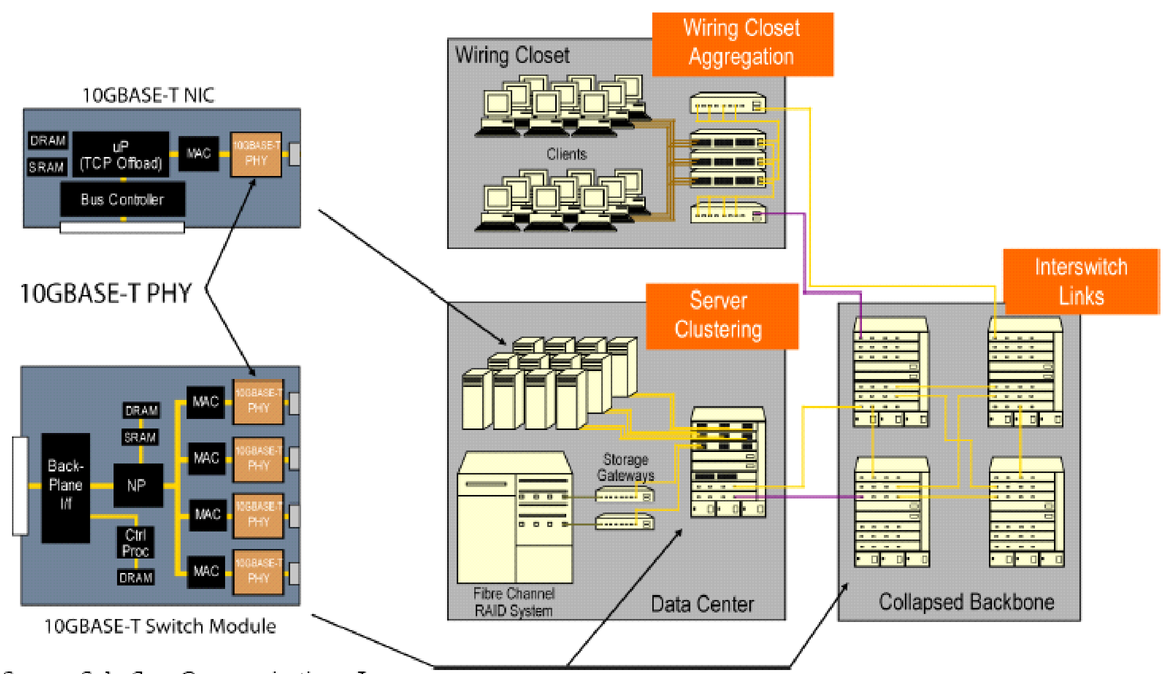
Figure 1—Interconnects Used on the Top 500 Supercomputers

3. Company and product names may be trademarks of their respective companies.

10GBASE-T Uplinks

One of the first applications for 10GBASE-T is to serve as an uplink technology offering high-bandwidth connectivity, from the access or distribution layer, to the next point of aggregation. Frequently such switch links are deployed over Category 5e or Category 6 cabling with 10/100/1000 Mb/s Ethernet desktop connections and a 1000BASE-T uplink. Switches with X2, XFP or XENPAK slots may be upgraded to 10GBASE-T uplinks, giving end stations the opportunity to upgrade from 1,000 Mbps links to 10,000 Mbps links and providing high-bandwidth, 10 Gb/s access to servers and other resources. There will also be fixed configuration switches with 10/100/1000 access ports and fixed 10GBASE-T uplink ports.

Figure 2 illustrates the expected early deployment of 10GBASE-T enabling higher bandwidth links in the data center and serving as an uplink technology offering high-bandwidth connectivity from desktop switches to the next point of aggregation.



Source: Solarflare Communications, Inc.

Figure 2—10GBASE-T Deployment

One PHY, Two Modes

From the time the 10GBASE-T project started in 2002, data centers were seen as one of the main target markets when 10GBASE-T solutions were delivered to the market in 2007. During 2005, the last year of the standards development effort, data center and server OEM customers communicated their desire for 10GBASE-T PHYs at power dissipation levels lower than some of the power estimates communicated during the standards process. During their November 2005 plenary meeting, the 802.3 Working Group, which charters all Ethernet projects including 10GBASE-T, rejected a proposal to develop a short-reach, reduced-power PHY. The main reasons for this rejection were concerns about the lack of interoperability with 100 meter 10GBASE-T PHYs and the economic folly of splitting the target data center market into two applications, long reach and short reach.

Throughout the 10GBASE-T project, both the server community and the cabling industry argued that 100 meter reach was the right solution to cover all data center applications, commercial enterprise applications and future desktop applications. In November 2005, the IEEE P802.3an (10GBASE-T) task force committed to investigating a solution that would provide reduced power consumption within the framework of the standards requirement of one PHY.

The PHY vendors participating in the standards process agreed that the project could maintain the requirement of one PHY by allowing implementations to support both 100 meter and shorter reach channels. Also, when supporting that shorter reach, the PHY could go into a lower power mode. Because power dissipation is implementation specific, IEEE 802.3 standards projects do not specify power dissipation, although often power dissipations trade-offs influence technology choices.

10GBASE-T and Power Dissipation

Early implementations of the PHY dissipate close to 10 watts. While in 1999, early implementations of 1000BASE-T single port PHYs consumed up to 7 watts, today a Gigabit Ethernet PHY consumes much less than 1 watt. The relatively higher power of 10GBASE-T PHYs can present a barrier to developing high-density switching systems or pluggable format solutions. 10GBASE-T technology offers three stages to resolve the problem.

First, as described above, the IEEE Std 802.3an-2006 standard includes a provision for a short reach mode, supporting 30 meters of Class E_A (Augmented Category 6) or Class F (Category 7) cabling. It is expected that the power requirements of the PHY will reduce by as much as 20 to 30 percent when operating in this mode. Some deployments of 10GBASE-T, in data centers or as inter-switch links, will use short reach, high- performance cable to benefit from this low power mode.

The power requirements of a 10GBASE-T PHY are not dominated by any one single component of the PHY. In the short term this causes a problem because each developer seems to be focused on reducing the power of one part of their chip, leaving the other parts to consume more power, but in the long term this encourages optimism that the overall PHY power will come down as knowledge is shared among developers. Later-generation designs will make use of more aggressive silicon technology and developers may feel free to take risks with novel techniques to achieve even lower power.

Data Center Type	Data Center Size (Sq. Ft.)	0-30 m	31-45 m	46-55 m	56-75 m	76-100 m
		Percentage of Cables				
Corporate	5000	100%	0%	0%	0%	0%
Corporate	8000	98%	2%	0%	0%	0%
Government	10000	70%	20%	10%	0%	0%
Corporate	20000	70%	25%	5%	0%	0%
Corporate	20000	90%	9%	1%	0%	0%
Internet	40000	60%	35%	5%	0%	0%
Corporate	45000	65%	25%	8%	1%	1%
Internet	60000	35%	48%	15%	1%	1%
Internet	60000	55%	35%	8%	1%	1%
Internet	80000	34%	56%	8%	1%	1%
Internet	100000	2%	98%	0%	0%	0%
Average		**61.7%**	**32.1%**	**5.5%**	**0.4%**	**0.4%**

Source: IEEE 802.3 10GBASE-T Study Group[4]

Table 1— Data Center Horizontal Cabling Lengths

4. DiMinico, Chris. MC Communications, "Data Center Design Considerations, "IEEE 802.3, http://www.ieee802.org/3/10GBT/public/nov03/diminico_1_1103.pdf, November 2003.

The Adoption of 10 Gigabit Ethernet and the Role of 10GBASE-T

Since its introduction in 2002, deployment of 10 Gigabit Ethernet networks has been relatively slow compared to previous variants of Ethernet. Historically, the common wisdom for justifying widespread adoption of the next speed of Ethernet has been achieving 10 times the throughput for three to four times the cost. In that respect, by 2007, 10 Gigabit Ethernet had not yet reached its full potential. The cost of deploying 10 Gigabit Ethernet is still much higher than the desired three times to four times the cost of Gigabit Ethernet. Also, at least on the end-node side, the throughput is much lower than true 10 Gb/s in many cases.

In order to speed up the adoption of this technology, the industry needs to tackle both the cost and the performance aspects of the problem. It is important to note that by performance we mean the actual usable networking throughput from an end user's perspective, and not the nominal speed of a link.

Cost

In terms of both Capital Expense (CapEx) for passive infrastructure and well as human operational expenses, 10GBASE-T promises to be a lower cost technology than optical alternatives because of its support of both legacy and new twisted-pair copper cabling. Almost all cabling inside buildings is copper, Category 5e or better. Therefore, 10GBASE-T products allow network planners to deploy 10 Gigabit Ethernet technologies over their installed cabling plants that meet the 10GBASE-T link segment specifications. The industry's support of installed horizontal cabling infrastructure with standards-based 10GBASE-T is crucial. While networking equipment can easily be pulled from a rack, horizontal cabling can be very difficult to replace since it is located inside a wall, ceiling, or raised floor and dispersed across many wiring closets.

As explained above, 10GBASE-T enables network managers to preserve their investment in existing Ethernet equipment. On a cost per port basis, 10GBASE-T ports cost less than optical 10 Gigabit Ethernet ports because copper transceivers are less expensive than optical transceivers. Today's cost per port of 10 Gigabit Ethernet implementations is still dominated by the physical layer devices (PHYs). Although the cost of these devices has been rapidly declining since the 10 Gigabit Ethernet standard was ratified in 2002, it still accounts for 40%-60% of the total cost per port.

10GBASE-T is expected to play a major role in continuing the above-mentioned trend for lowering the cost of 10 Gigabit Ethernet implementations. Its low initial cost, combined with the traditional advances in silicon technologies (Moore's Law), will position 10 Gigabit Ethernet on a similar declining cost curve that Ethernet users have become accustomed to.

Although traditionally Ethernet technology is a high cost for early adopters, as the volume of ports shipped increases, costs come down, eventually reaching commodity pricing. In some cases the initial cost-performance ratio can be as low as a sevenfold increase in bandwidth for 10 times the cost.[5] The performance-cost ratio that drives Ethernet towards commodity pricing is a 10X bandwidth increase for a 3X cost increase as is evident in the growth of Ethernet over competing technologies. For 10 Gigabit Ethernet, 10GBASE-CX4 has achieved the expected Ethernet performance-cost ratio, however this technology has installation limitations similar to non-Ethernet interconnects.

Performance

One of the reasons for the relatively slow adoption of 10 Gigabit Ethernet on end-nodes in the data center (servers) is that, at the beginning of 2006, a typical network interface card (NIC) was not capable of sustaining a throughput that is anywhere close to the nominal rate of 10 Gb/s. This drives the deployment cost of 10 Gigabit Ethernet (per gigabit of usable throughput) even higher. Therefore, one of the most common questions raised by Ethernet users considering deployment of 10 Gigabit Ethernet in the data center is, "What is the ‘killer application’ for this new technology?"

Rather than one new application that drives the adoption of multi-gigabit networking on servers, a vast majority of existing network-facing applications already require this network bandwidth. This conclusion is based on the following observations:

- Since Gigabit Ethernet was introduced in the late 1990's, CPU speeds have increased by at least a factor of 4, and in some cases by a factor of 8.
- During the same period, system memory speeds have doubled and, in some cases, quadrupled.[6]

5. The current host interface bus limits prevent achieving 10 Gb/s throughput.

6. Muller, S., SUN Microsystems, IEEE Higher Speed Study Group, 2007.

Thus, the same network-facing applications are now running much faster, and are often limited by the network throughput. In addition, the following trends in general purpose computing accelerate the need and the adoption of high-speed networking in servers:

- **Chip Multi-Threading (CMT).** Modern microprocessor architectures are based on multiple CPU cores on the same die, with each core having multiple hardware threads. This allows for both faster execution of multithreaded applications and, most importantly, running multiple applications on the same processor.
- **Server Consolidation.** The use of CMT in servers allows users to consolidate multiple server boxes that are running multiple applications with very low CPU utilization, into one server box that does the same amount of work with greater efficiency.
- **Virtualization Technologies.** These technologies enable the abstraction of the server resources into a common pool, and partition them on an as needed basis for each application. Today these techniques are commonly used for compute elements, memory and storage. In the near future they will also be applied to general purpose I/O and networking.

All of the above will allow an intelligent 10 Gigabit Ethernet NIC to be shared among multiple applications running on multiple threads.

From a purely server networking throughput perspective, the above trends significantly simplify the solution space. The new paradigm described above, moves the first traffic aggregation point from the switch to the server NIC. Therefore, the optimization point in a NIC implementation no longer needs to be a speed up of a single high-speed network connection, but rather scaling the throughput of multiple lower-speed connections that run on multiple threads. And that is a much simpler problem to solve without resorting to complex and expensive solutions, such as offload engines. Moreover, as the industry approaches the end of 2007, and as host CPUs get faster and move from dual to quad to octal cores, server adapter I/O will improve as well without the need for such off load engines.

Latency

The 10GBASE-T PHY is specified with a round-trip delay of up to ~2.5 microseconds. Although this can result in an absolute latency substantially higher than what it used to be for the previous variants of Ethernet, it is not expected to affect application performance for the vast majority of environments where Ethernet is deployed today and 10 Gigabit Ethernet is expected to be deployed in the future.

How 10GBASE-T Works: The Technology

This section explores the technology behind 10GBASE-T.

10GBASE-T Signaling: An Overview

The migration from lower data-rate twisted-pair Ethernet to higher rates [e.g., Fast Ethernet (100BASE-T) to Gigabit Ethernet (1000BASE-T), or Gigabit Ethernet (1000BASE-T) to 10 Gigabit Ethernet (10GBASE-T)] involves a further exploitation of the available bandwidth on the twisted-pair copper cabling. In order to accomplish this, each generation of Ethernet uses both more efficient signaling, and increased signal processing mitigation of the impairments on the cabling.

In 1000BASE-T, Ethernet evolved to using all 4-pairs in a Category 5e cable, multi-bit per baud trellis-coded modulation, and bi-directional signaling (requiring both echo and near-end crosstalk (NEXT) cancellation on each pair). In order to get another order of magnitude increase in the bit rate, 10GBASE-T takes this several steps further, increasing the signaling rate (from 125 Mbaud to 800 Mbaud), the number of levels in the pulse amplitude modulation (PAM) transmitted signal (from 5 to 16 levels), and the performance of the error correcting code (by employing state-of-the-art low density parity check (LDPC) codes). As well it substantially improves receiver sensitivity, echo and crosstalk cancellation. It is these improvements in receiver sensitivity and interference cancellation which make 10GBASE-T possible; however, in order to operate efficiently at higher speed, several technology advances have been taken advantage of in 10GBASE-T. Table 2 below summarizes the key technical elements of 1000BASE-T and 10GBASE-T.

1000BASE-T	10GBASE-T
5-level coded PAM signaling (2 information bits/symbol)	Baseband 16-level PAM signaling (3.125 information bits/symbol)
8-state 4D Trellis code across pairs	128-DSQ + LDPC(2048,1723) Tomlinson-Harashima precoder at Tx
Full duplex echo-cancelled transmission	Full duplex echo-cancelled transmission
125 Mbaud, ~ 80 MHz used bandwidth	800 Mbaud, < 500 MHz used bandwidth
FEXT Cancellation recommended	FEXT Cancellation required

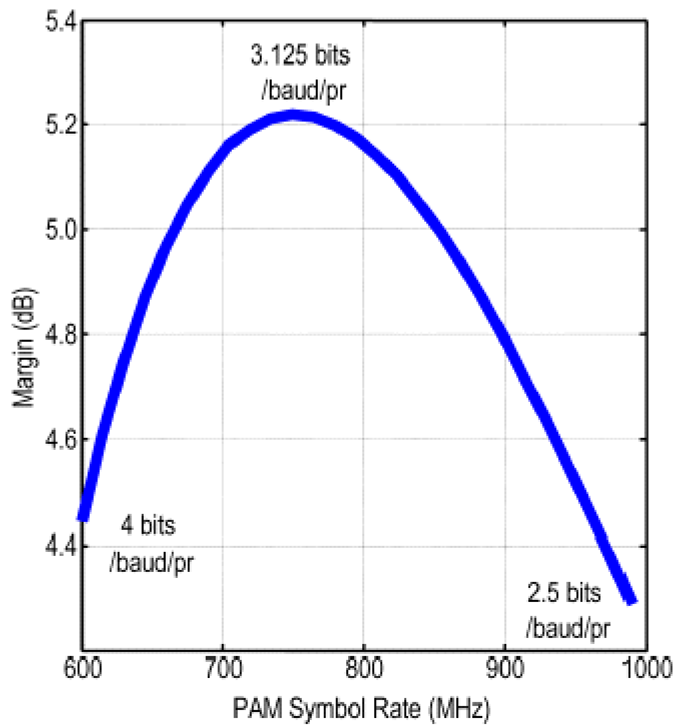
Source: Solarflare Communications, Inc.

Table 2 – Summary of Key Technical Elements and Comparisons

Because the twisted-pair cabling channel attenuates increasingly with frequency, the bandwidth which can be used on a cable is a function of the distance and the residual noise level. The usable bandwidth is limited by the point at which the uncanceled noise received at and above a frequency is generally greater than the signal received from the far end at those same frequencies.

Previous generations of Ethernet were limited by internal NEXT which crossed-over with the attenuation at a frequency between 100 and 250 MHz (depending on the cable category). 10GBASE-T, however, cancels the internal NEXT to levels where neither it, nor any other internal pair-to-pair crosstalk are the dominant impairment. Instead, alien crosstalk, from other cables in close proximity limits 10GBASE-T performance. The limitations from alien crosstalk and externally-induced background noise put the usable bandwidth near 500 MHz, with lower frequencies having higher signal-to-noise ratios (SNR) than higher frequencies.

A balancing act is set up between achieving higher SNR by using less bandwidth (hence more bits/symbol) in the signaling, and the consequence that this requires greater SNR to support the higher number of bits per symbol. A common way to express this is with "SNR margin" - the amount by which the SNR obtained exceeds that required to maintain the desired error rate. The IEEE 802.3 10GBASE-T study group and the IEEE P802.3an task force undertook studies to determine the optimum signaling rate on the cabling link segment defined for 10GBASE-T, and found it to be near a PAM symbol rate of 800 MHz, as shown in Figure 3.



Source: IEEE P802.3an 10GBASE-T Task Force

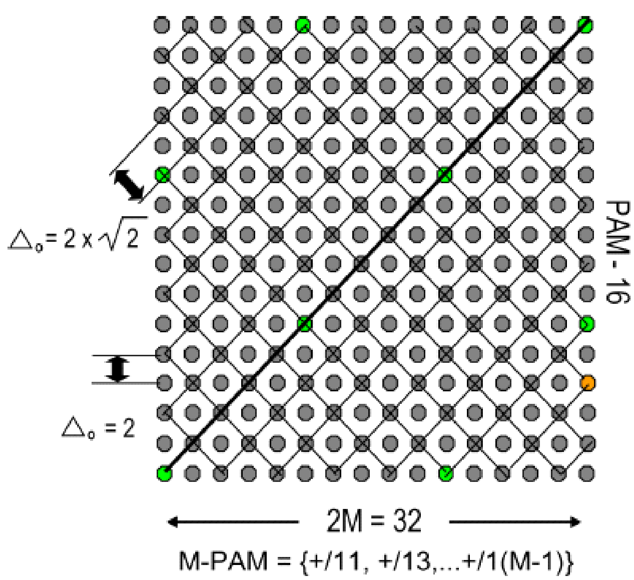
Figure 3— SNR Margin vs. Symbol Rate

Constellation and Encoding

Transmitting 10 Gb/s over four pairs at a PAM symbol rate of 800 MHz means that on each twisted-pair, 3.125 information bits need to be transmitted during each baud interval. Since information bits are the bits received after accounting for forward error correction (FEC) overhead, translating this to a specific signal level requires knowledge of the code overhead and coding structure used. The constellation is a diagram used to represent the allowed signaling levels.

For 10GBASE-T, a very high-performance class of FEC codes was used. These codes are known as low density parity check, or LDPC, codes. LDPC codes were discovered in 1960 by Gallagher, and are capable of achieving performance near channel capacity. They were recently rediscovered in 1995 by MacKay and Neal. This rediscovery occurred due to advances in computing and integration capabilities in the decades since Gallagher's initial discovery. Such advances enabled high-performance systems to be economically built with these codes using iterative techniques. For 10GBASE-T, a low-overhead, high-performance (2048, 1723) LDPC code was constructed, which was shown to achieve 1e-12 bit error ratio (BER) with minimum SNR.

In order to not use too many PAM levels, the code protects only a subset of the levels in the PAM signal, similar to 1000BASE-T, using techniques pioneered by Ungerboeck. However, while in 1000BASE-T the code protected only the nearest two levels, creating a 6 dB partition (for 6 dB maximum coding gain); the more powerful LDPC code in 10GBASE-T protects subsets of four levels for a 12 dB partition. Including the low-overhead code and using an efficient constellation partitioning and some minimal framing ultimately required the constellation to map a total of 3.5 coded and uncoded bits to each 800 Mbaud symbol on each of the four twisted pairs. This is accomplished by grouping two symbols, adjacent in time, together into a 7 bit word, and mapping them to a 128 point two-dimensional (one for each time slot) constellation known as DSQ128 (for Double-Square, 128 points). The DSQ128 symbols are thus determined by 7-bit labels, each comprising three uncoded bits and four LDPC encoded bits. DSQ128 is formed by sending one of 16-levels of pulse amplitude modulation (PAM16) at each time interval on each pair in such a way that when the allowed levels sent in two adjacent time intervals are considered two-dimensionally, they form a checkerboard pattern, as shown in Figure 4.



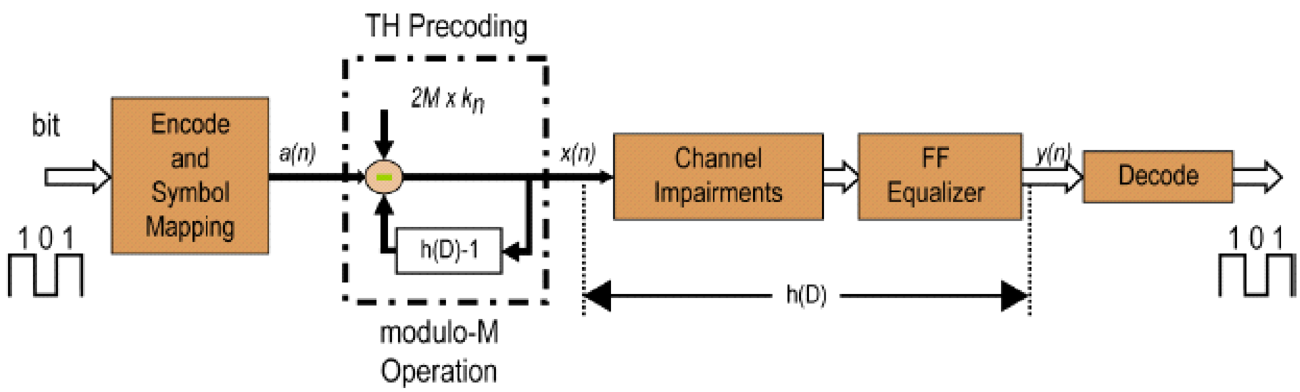
Source: IEEE P802.3an
10GBASE-T Task Force

Figure 4 – DSQ128 Constellation

Tomlinson-Harashima Precoder

The final step in preparing the signal for transmission is to pre-equalize it. This is done using a device known as a Tomlinson-Harashima Precoder, or THP. The purpose of a THP is similar to a decision feedback equalizer (DFE), fairly common in modern equalized systems and used in earlier generations of Ethernet (e.g., 1000BASE-T). The primary purpose of a DFE is to allow some of the filtering necessary for equalization to operate only on the received signal, and therefore avoid enhancing the noise received on the channel. In a DFE this is done using the decided symbols through a feedback-form equalizer filter. This becomes difficult when a very powerful code is used, such as the LDPC code in 10GBASE-T, since individual decisions contain a large percentage of errors. The result can be a catastrophic buildup of errors in a receiver, known as error propagation, and can make the straightforward implementation of a DFE useless.

Fortunately, this problem has been solved before in both modem design and in digital subscriber line (DSL) transceivers. The THP solves this problem by effectively moving the DFE into the transmitter, where the transmitted symbols are known. The feedback filter is not necessarily stable. To compensate for this, the THP adds, when necessary, an additional quantity to the signal to keep the amplitude at the transmitter bounded. This quantity is designed so that it does not interfere with decoding the bits used in the constellation at the receiver. While this sounds conceptually like a hard problem, the solution turns out to be fairly simple. By performing a modulo operation, which can be viewed as adding (or subtracting) a multiple of the full-scale range of the transmitter to the symbol to be transmitted, the transmitted signal is kept within a limited range (-L, L]. The receiver then employs an *expanded slicer* which detects an expanded symbol ck = ak + 2L.ik, where L is the order of the symbol alphabet, ak is the transmitted symbol, and ik is an integer chosen to limit the transmitter output to the range (-L,L]. The result is that the transmitted symbol (ak) is equalized and easily recovered, and peak voltages at the transmitter are kept within a minimal range. The main hardware cost of Tomlinson-Harashima precoding is largely born by the transmitter digital-to-analog conversion, which now must output a uniformly-distributed smear with an accuracy approaching 10 bits, even though the transmitted symbol before the precoder only has 16 discrete values.



Source: Solarflare Communications, Inc.

Figure 5 – Tomlinson-Harashima Precoder

PHY Training and Startup Sequence

The Tomlinson-Harashima precoder coefficients must be exchanged between the two link partners during startup. For this reason, 10GBASE-T transceivers must have a well-codified, sequential startup sequence. The participants writing the IEEE Std 802.3an-2006 Standard knew that the robustness and specificity of this startup sequence would be important for the success of 10GBASE-T transceivers. 10GBASE-T startup builds on prior generations of Ethernet. The startup sequence initiates with the standard twisted-pair Ethernet auto-negotiation protocol. Transceivers exchange capabilities and agree on the highest speed that both can support. While loop timing is not mandatory in 10GBASE-T, transceivers also resolve which will be the loop-timing MASTER and which will be the SLAVE, similar to 1000BASE-T. If loop timing is not supported by one of the devices, it defaults to the MASTER. In this way conflicts are avoided and PHY designers have flexibility in timing recovery trade-offs.

After exiting auto-negotiation, the 10GBASE-T PHY enters training mode. Training begins with the SLAVE silent and the MASTER at a reduced transmit power until it receives a response from the slave. 10GBASE-T PHY training is coordinated by exchanges of information using 2-level PAM signaling in "Infofields". Infofields are used throughout 10GBASE-T training to convey state transitions, transmit power settings, transition synchronization information, receiver status (e.g., SNR), and exchange of precoder coefficients.

New to Ethernet, the advanced coding in 10GBASE-T allows it to test the link's data integrity at the PHY level. As the last stage of the startup sequence, the transceivers switch to transmitting the 128-DSQ signaling and enable the LDPC coding, and test that the link is actually receiving good frames.

Link Quality Monitoring

In addition to this initial test and other frame-error measurements, the 10GBASE-T standard includes optional registers that are intended to provide a report of the SNR margin, a measure of the quality of the physical layer before errors begin. The SNR margin is relative to the SNR required for the reception of LDPC-coded DSQ128. Registers are assigned to each of the cable pairs and the current SNR operating margin measured at the slicer (receiver) input for each pair can be reported.

The SNR margin variable may be applied to determine if the link has sufficient operational SNR margin. Understanding operational margin may be useful when deploying 10GBASE-T over cabling types and distances which are not explicitly supported.

Power Backoff

Since the advanced signal processing in 10GBASE-T manages to mitigate the sources of noise and distortion within an individual 4-pair cable sheath, this ultimately leaves the system limited in performance by external crosstalk from other cables in close proximity, which is called "alien" crosstalk. Alien crosstalk can come either from sources at the same end of a transmitting link (alien NEXT) or can come from sources which transmit at the other end, or somewhere along the length of the link, and couple down the length of the link (alien FEXT).

Alien NEXT sources can be effectively mitigated through cabling design, qualification, and installation practices as described in other sections of this white paper; however, alien FEXT proved to be a tougher problem for cabling. This is due to the fact that the disturbing far-end transmitter could be much nearer to the disturbed receiver than that receiver's own intended link partner. Fortunately this problem, known as the "near-far problem" has been addressed before in both wireless and wireline (DSL) systems. To avoid the near-far problem, transmitters are required to reduce their transmit power to only the level needed on their link.

Thus, short link transmitters will transmit with a "power backoff" from the nominal power used for longest, 100 meter links. Power backoff is not to be confused with a lower power mode of PHY operation.

Figure 6 illustrates two link segments in close proximity with horizontal cable lengths of 15 meters and 90 meters respectively. The transmit power of Link 1 is backed-off by 10 dB which reduces the alien crosstalk power coupled into Link 2 by 10 dB. The power backoff (PBO) level of 10 dB on Link 1 is allowed because the receiver power level of Link 1 is within the specified limits due to the relatively short link length of 15 meters.

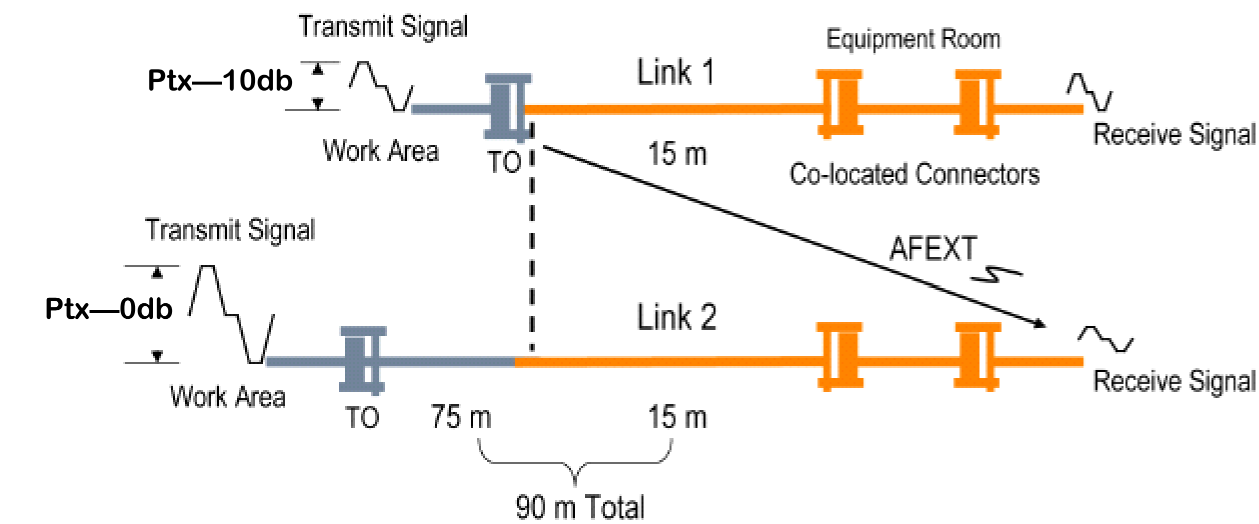


Figure 6 – Power Backoff

The Link 2 PBO level allowed is 0 dB due to the link length of 100 meters. When existing links are qualified for 10GBASE-T computation, power backoff considerations are included in the alien crosstalk computation.

All standard compliant 10GBASE-T transceivers will follow a minimum power backoff scale which was carefully designed to ensure that networks maintain robust performance. It is important to note that introducing additional wide-spectrum signals into the unshielded twisted-pair (UTP) network, such as proprietary PHY technologies utilizing signaling rates greater than 125 MHz, may create interference problems for both these signals and standards-compliant

10GBASE-T. Thus, care should be taken for compatibility. 10GBASE-T has been designed so that existing lower rate Ethernet systems (e.g., 1000BASE-T) which do not have power backoff are compatible. This was possible because these legacy systems use much less signaling bandwidth than 10GBASE-T.

How You Use 10GBASE-T: Cabling and Field-testing Issues

10GBASE-T Cabling Considerations

The 10GBASE-T link segment specifications are the minimum cabling requirements specified to support 10GBASE-T operation. The transmission parameters of the link segment include insertion loss, return loss, pair-to-pair NEXT Loss, power sum NEXT Loss, pair-to-pair ELFEXT, power sum ELFEXT, return loss, and delay. The link segment specifications are based on the Class E channel limits[7] extended to an upper frequency of 500 MHz.

In addition, alien crosstalk, crosstalk coupled "between" link segments, is specified for power sum alien NEXT (PSANEXT) and power sum alien ELFEXT (PSAELFEXT). The link segment cabling topology, consisting of components (cables, cords, and connectors), is illustrated in Figure 7.

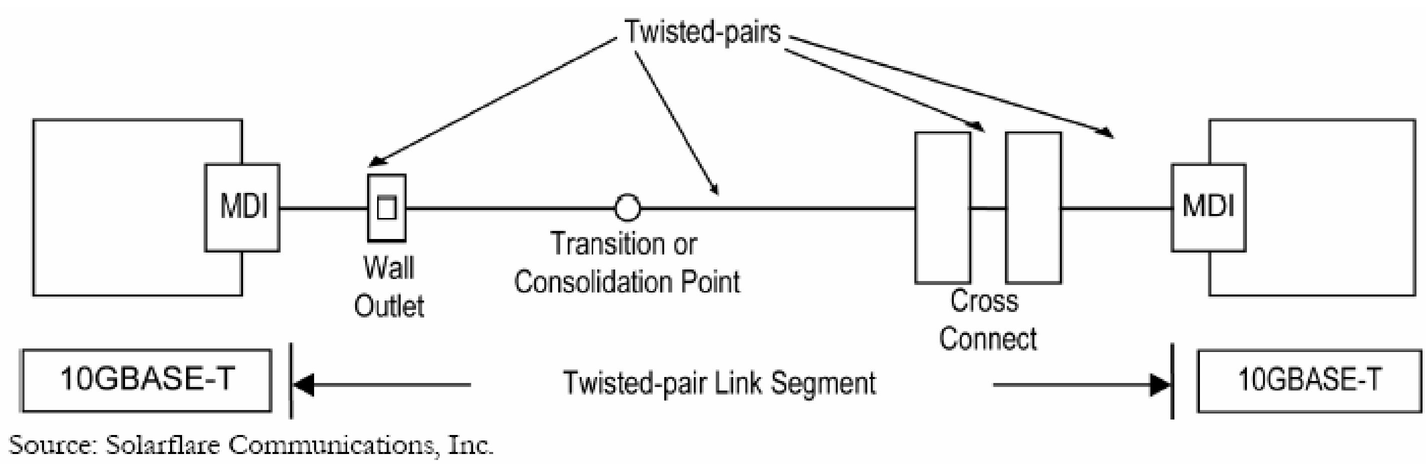


Figure 7—Twisted-pair Link Segment

7. The ISO/IEC 11801 Class E cabling channel specification are approximately the same as the ANSI/TIA-568B.2-1 Category 6 cabling channel specification.

10GBASE-T Cabling Types and Distances

The 10GBASE-T cabling distances supported depend on the insertion loss and the alien crosstalk performance of the cabling. For Class E cable constructed with a foil screen (FTP) with suitable screened connecting hardware, Augmented Category 6 cabling, and Class F cabling, the measured alien crosstalk performance is better than the 10GBASE-T alien crosstalk requirements. Thus, lengths of 100 meters can be supported. FTP is referred to as screened twisted-pair (ScTP) in TIA-568-B.1. For Class E UTP, the 55 meter limit was established based on alien crosstalk measurements of installed UTP cabling representative of a "reasonable" worst case.

Operation on other classes of cabling (for example, Class D/Category 5e) is supported if the cabling meets the 10GBASE-T link segment specifications or the alien crosstalk margin computation. The 10GBASE-T link segment specifications satisfy the minimum requirements of the cabling types and distances identified in the objectives. 10GBASE-T supports operation over:

- At least 100 meters on 4-pair Class F balanced copper cabling; and
- At least 55 meters to 100 meters on 4-pair Class E balanced copper cabling.

Cabling	Support Link Segments Distances	Cabling Reference
Class E / Category 6	55 to 100 meters	ISO/IEC TR-24750 / TIA/EIA TSB-155a
Class E / Category 6 Unscreened	55 meters	ISO/IEC TR-24750 / TIA/EIA TSB-155
Class E /Category 6 Screened	100 meters	ISO/IEC TR-24750 / TIA/EIA TSB-155
Class F	100 meters	ISO/IEC TR-24750
Class EA /Augmented Category 6	100 meters	ISO/IEC 11801 Ed 2.1 /TIA/EIA-568-B.2-10b

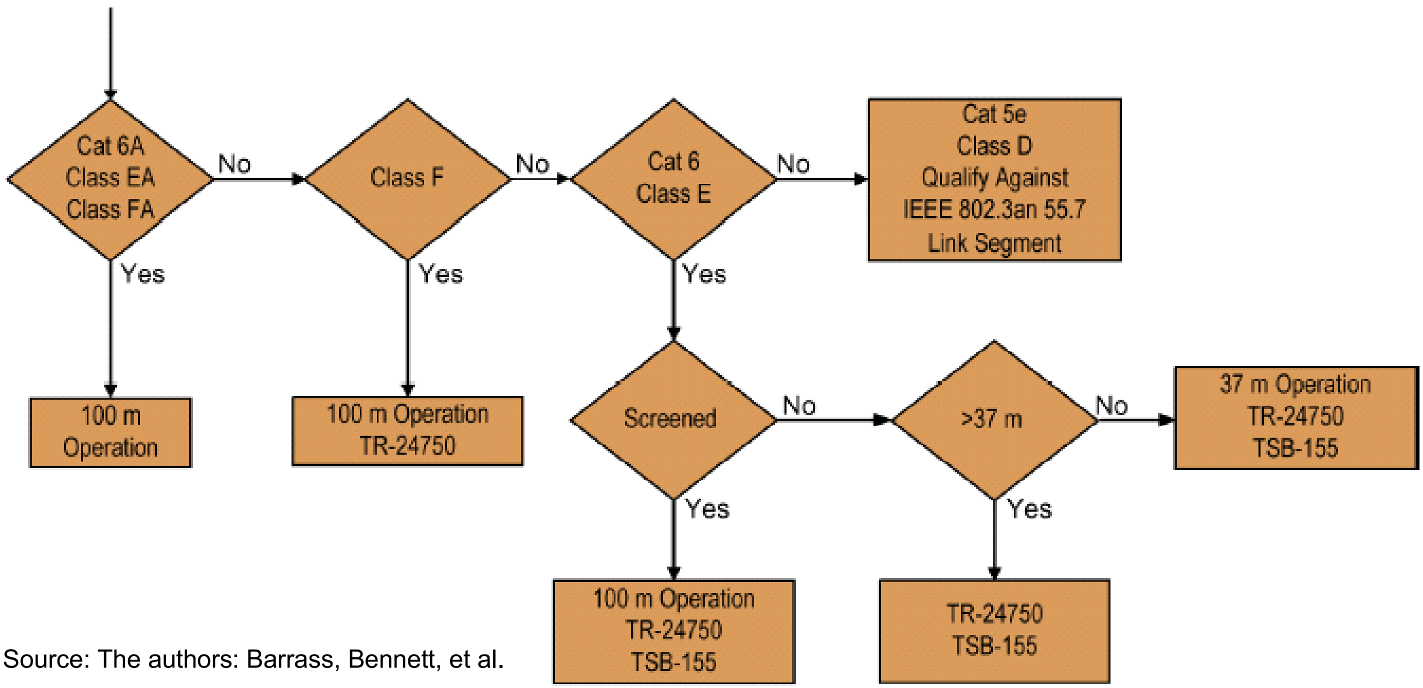
Source: IEEE 802.3an 10GBASE-T Task Force

a. Additional Cabling Guidelines for 4-pair 100 Category 6 Ω Cabling for 10GBASE-T detailed in section “Appendix A: TR42-10GBASE-T Projects” on page 19.

b. Transmission Performance Specifications for 4-pair Ω 100 Augmented Category 6 Cabling detailed in section “Appendix A: TR42-10GBASE-T Projects” on page 19.

Table 3 – IEEE 802.3an 10GBASE-T Cabling Types and Distances

The flow diagram provided in Figure 8 illustrates the supported cabling types, supported link segment distances and cabling standard references.



Source: The authors: Barrass, Bennett, et al.

Figure 8 – 10GBASE-T Cabling Selection Flow Diagram

The industry tracking firm, BSRIA, has published data about the historical and projected installed base of structured cabling by outlet represented in Figure 9. BSRIA labels as 10G all the cabling types that support 10GBASE-T at 100 meters: Category 6 ScTP, Category 6A UTP and ScTP, and Category 7. The data shows that a significant fraction of the installed base has already installed Category 6 ScTP, Category 6A, and Category 7. The forecast also shows that installation of Category 5e has ceased and that the deployment of Category 6 will rapidly shift to the higher performance cabling types, primarily Category 6A.

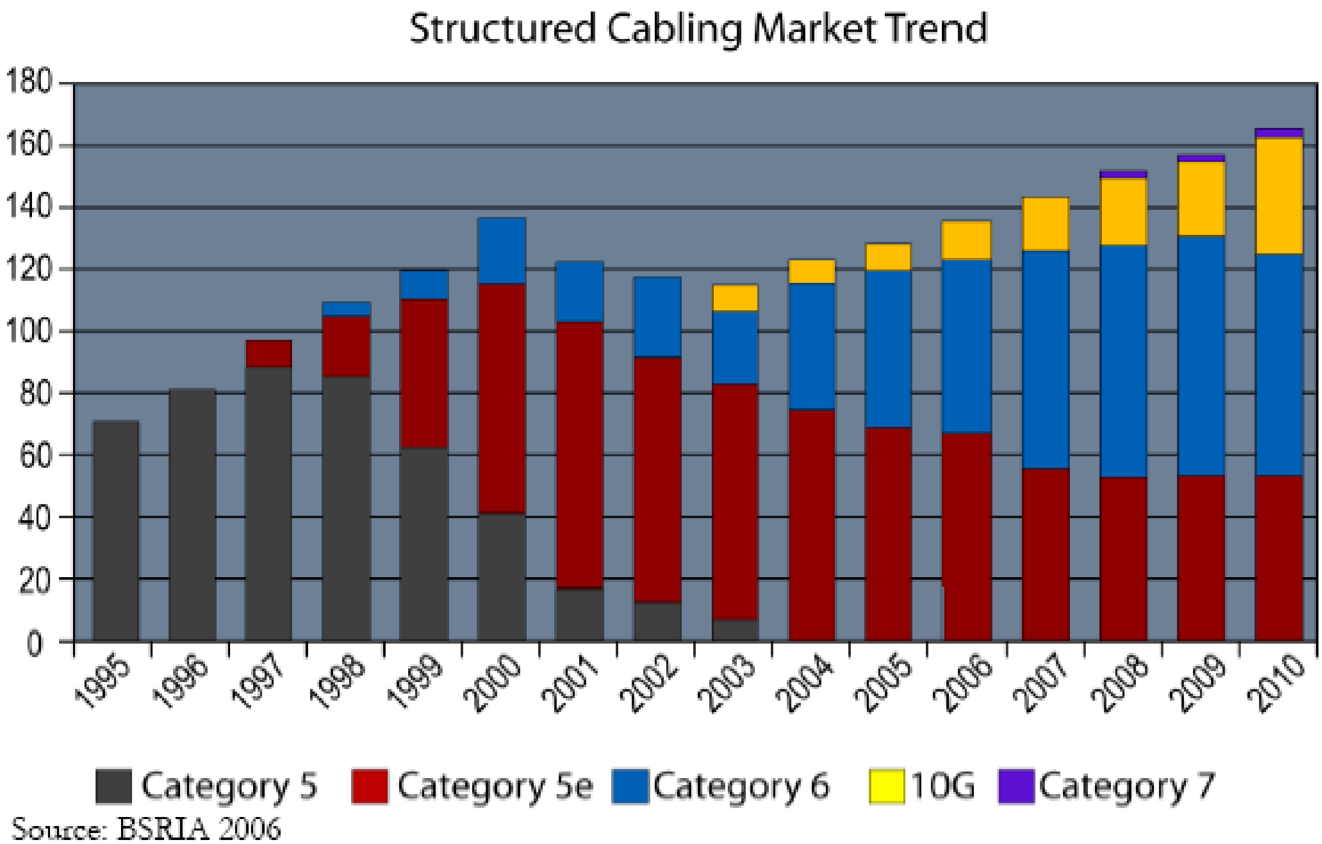
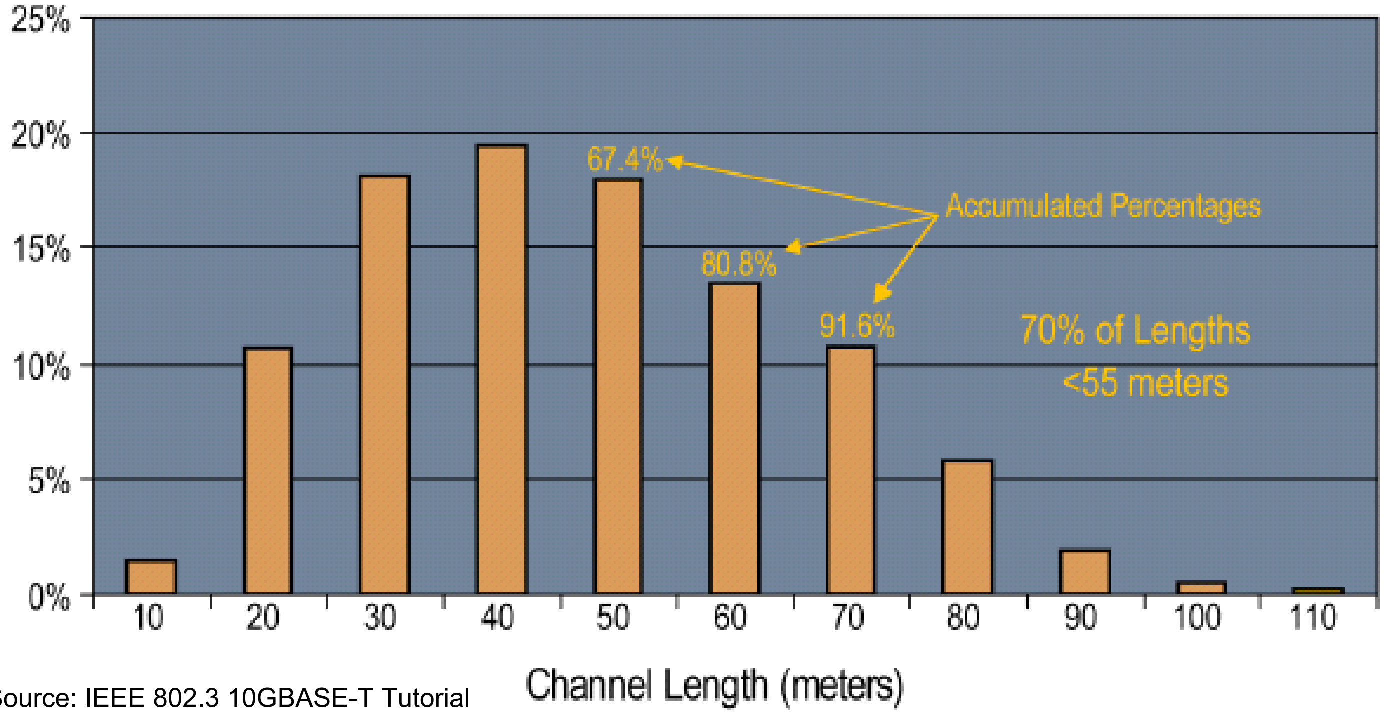


Figure 9 – Global Market Installation History and Trend by Category

Why 55 Meters on Class E UTP?

The IEEE P802.3an Task Force established a minimum operating range objective to ensure 10GBASE-T operation over a reasonable percentage of the Class E UTP installed cabling. The working group converged on a minimum operating range of 55 meters based on contributions from multiple sources. Figure 10 shows the percentages of installed cabling channels versus channel length contributed by several cabling manufacturers illustrating that for the lengths investigated 70 percent were less than 55 meters.



Source: IEEE 802.3 10GBASE-T Tutorial

Figure 10 – Horizontal Cabling Distances

Alien Crosstalk

Alien crosstalk is defined as crosstalk from cables adjacent to the cable that makes up the transmission channel. While other sources of noise such as near-end crosstalk and far-end crosstalk are mitigated by the PHY itself, for 10GBASE-T alien crosstalk is mitigated by the cabling media. Therefore, in reference to cabling, alien crosstalk and insertion loss are the primary parameters to consider for 10GBASE-T operation. In addition to the cabling transmission and coupling parameters, alien crosstalk as a function of insertion loss is specified to enable the evaluation of the combined affect of alien crosstalk and insertion loss. Alien crosstalk is specified as power sum alien near end crosstalk (PSANEXT), power sum alien equal level far end crosstalk (PSAELFEXT), and power sum alien far-end crosstalk (PSAFEXT). The 10GBASE-T SNR is primarily based on the ratio of the signal (insertion loss) to the alien crosstalk noise appearing at the receiver. The cabling impairments of return loss (echo), NEXT and FEXT within a cable are reduced to a small residual noise utilizing cancellation methods. Figure 11 illustrates the cabling impairment cancellation stages.

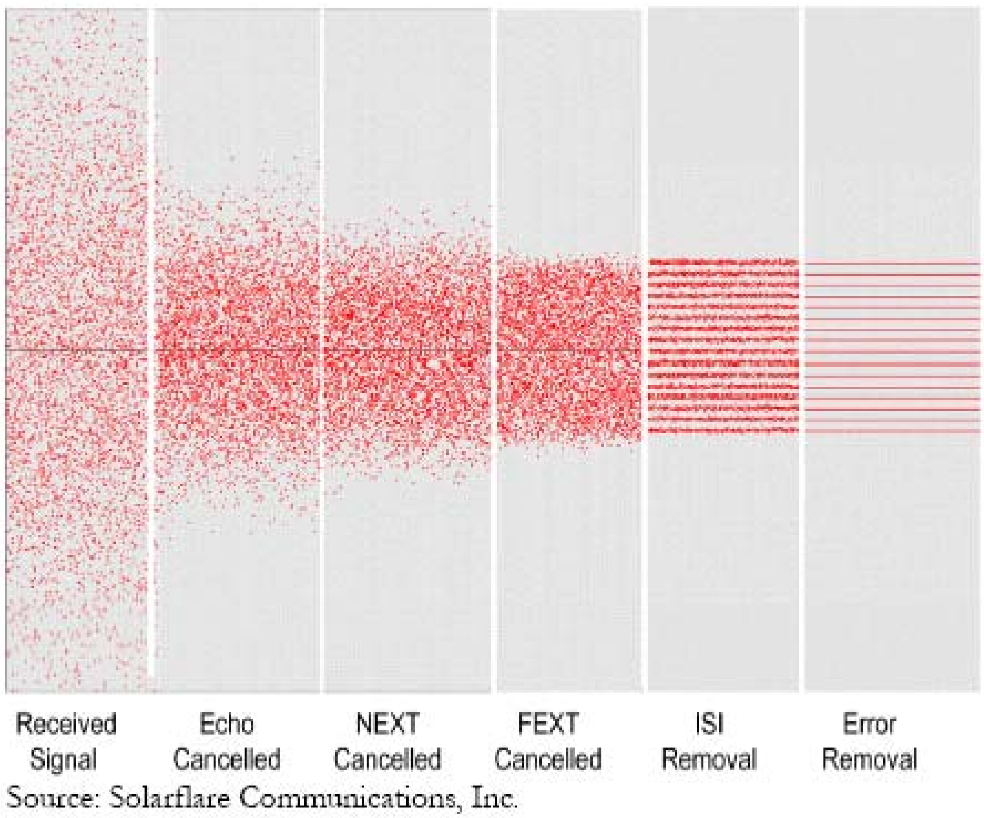
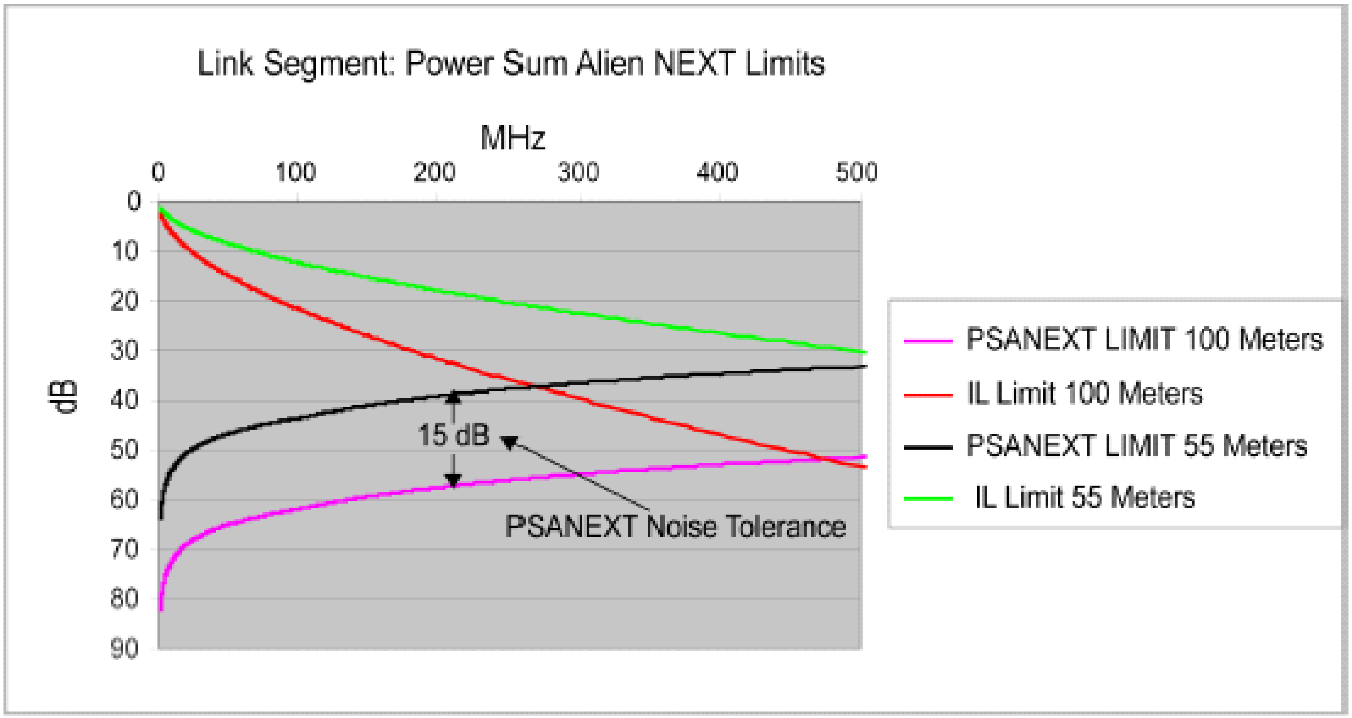


Figure 11 – Cabling Impairment Cancellation

The receive signal is initially indistinguishable from the noise and subsequently completely recovered after cancellation and intersymbol interference (ISI) correction.

The 10GBASE-T standard includes equations to determine the amount of PSANEXT loss and PSAELFEXT that a receiver can tolerate based on the link segment insertion loss. In addition, the average PSANEXT loss and average PSAELFEXT "across the 4-pairs" is specified to account for the benefits of the channel coding effectively averaging the noise "across all 4- pairs". The derived PSANEXT loss and PSAELFEXT noise tolerances and the insertion loss are applied as limits in qualifying installed cabling as well as specifying new cabling.

Illustrated in Figure 12, a link segment of 55 meters can tolerate 15 dB more PSANEXT noise than a link segment of 100 meters.



Source: Solarflare Communications, Inc.

Figure 12—10GBASE-T-PSANEXT Limites

Similarly, the PSAELFEXT loss limits determined from the PSAELFEXT to insertion loss requirements result in a 4.22 dB PSAELFEXT noise tolerance for the 55 meter link segment as compared to the 100 meter link segment.

Alien Crosstalk Margin Computation

Link segments that can support 10GBASE-T operation (i.e., those that have sufficient SNR margin) can still fail the individual pair limit lines. Figure 13 illustrates PSAELFEXT and PSANEXT measured on Pair 2 of a link segment.

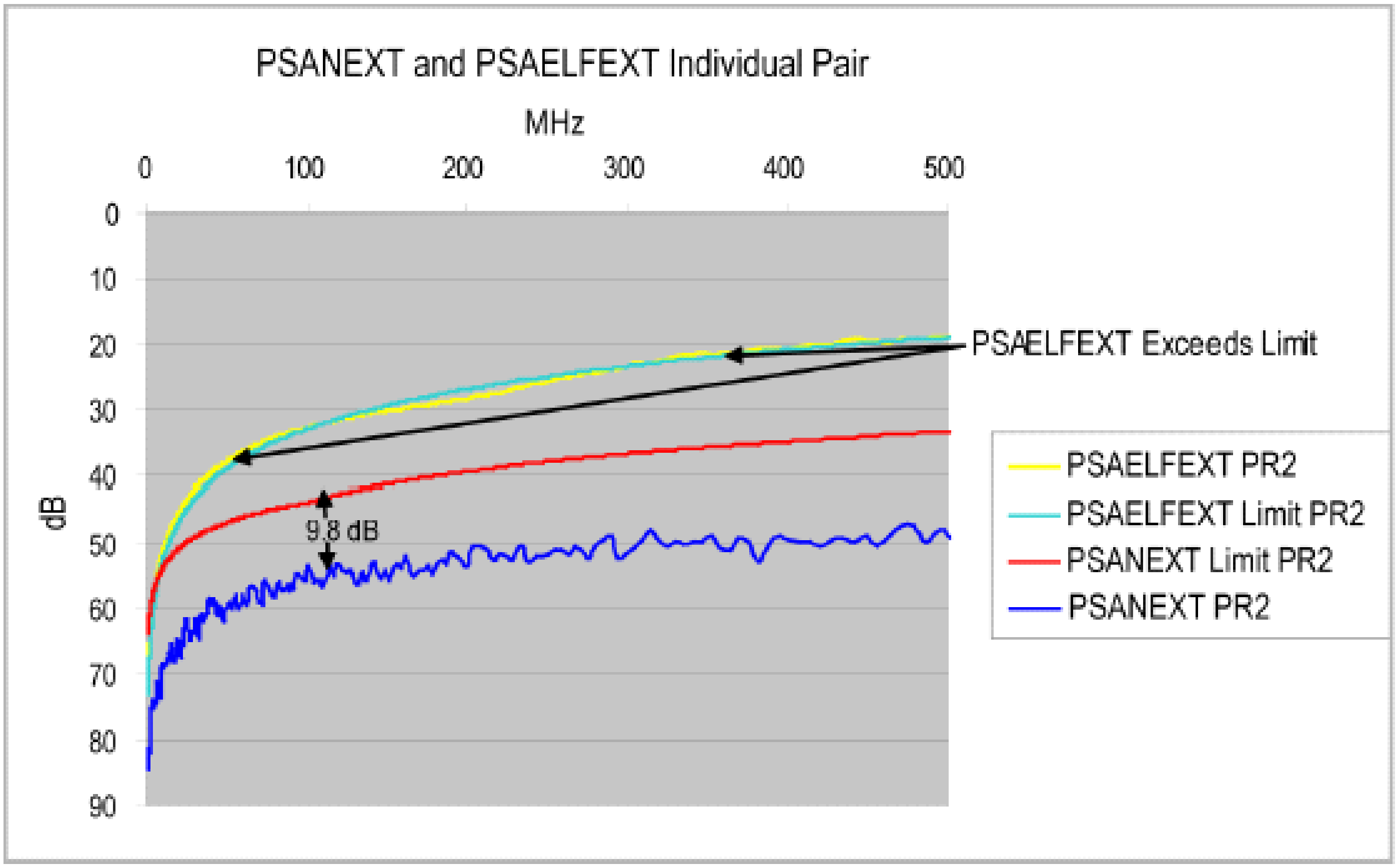
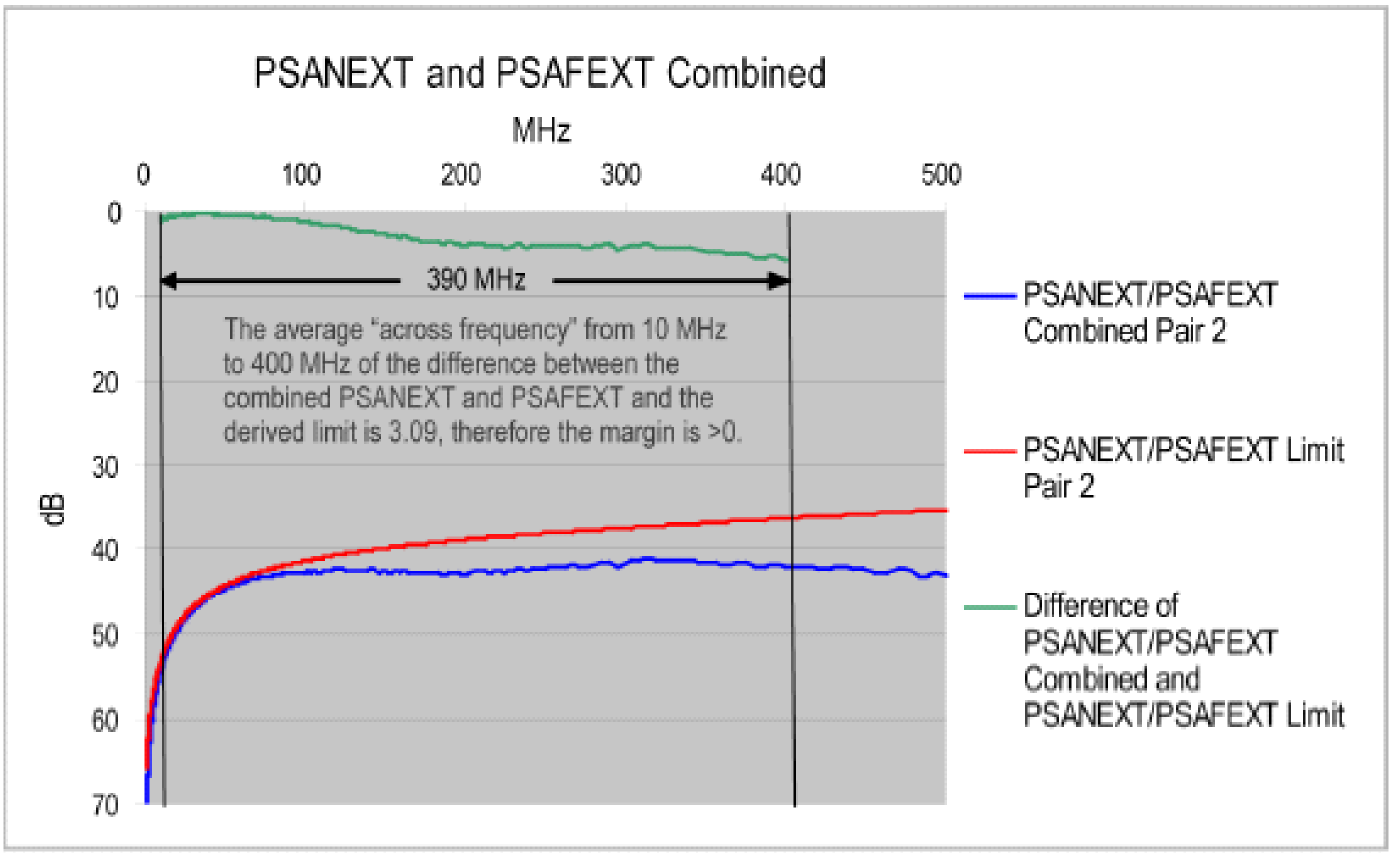


Figure 13 – Individual Pair Measurements

The PSAELFEXT fails the individual limit while the PSANEXT exhibits approximately 10 dB margin from the limit. The overall SNR margin for this link is sufficient to support 10GBASE-T operation because the 10GBASE-T receiver function does not discriminate between the individual PSAELFEXT and PSANEXT noise contributions. The receiver detects symbols from the "combined noise" and the transmitted signals on each of the 4-pairs.

10GBASE-T specifies an alien crosstalk margin computation to assess whether the cabling can support 10GBASE-T operation in the event that the individual pair limits are not met. The combined PSANEXT and PSAFEXT noise is calculated and subtracted from the combined alien crosstalk noise limit determined from the disturbed pairs insertion loss. The average "across frequency" from 10 MHz to 400 MHz of this difference is the alien crosstalk margin. The margin must be greater than zero to utilize the cabling for 10GBASE-T operation. The alien crosstalk margin is specified for each of the individual 4-pairs as well as the average "across the 4-pairs."

Figure 14 illustrates the combined PSANEXT and PSAFEXT and the calculated margin of Pair 2 which had failed the individual pair limits (Figure 13). Although the pair had failed the PSAELFEXT limit, the alien crosstalk margin computation ensures that the total combined PSANEXT and PSAFEXT is limited to maintain the minimum signal to noise ratio and therefore Pair 2 supports 10GBASE-T operation.



Source: Solarflare Communications, Inc.

Figure 14 – PSANEXT and PSAFEXT Combined

Alien Crosstalk Mitigation

Additional cabling guidelines for 10GBASE-T deployment on balanced copper cabling are provided in an informative Annex (Annex 55B) published in the 10GBASE-T standard. Annex 55B includes procedures to mitigate alien crosstalk in the event that the link segment specifications are not met ANSI/TIA/EIA-TSB-155 "The Additional Cabling Guidelines for 4-pair 100 Ω Category 6 Cabling for 10GBASE-T Applications" also provides procedures for alien crosstalk mitigation.

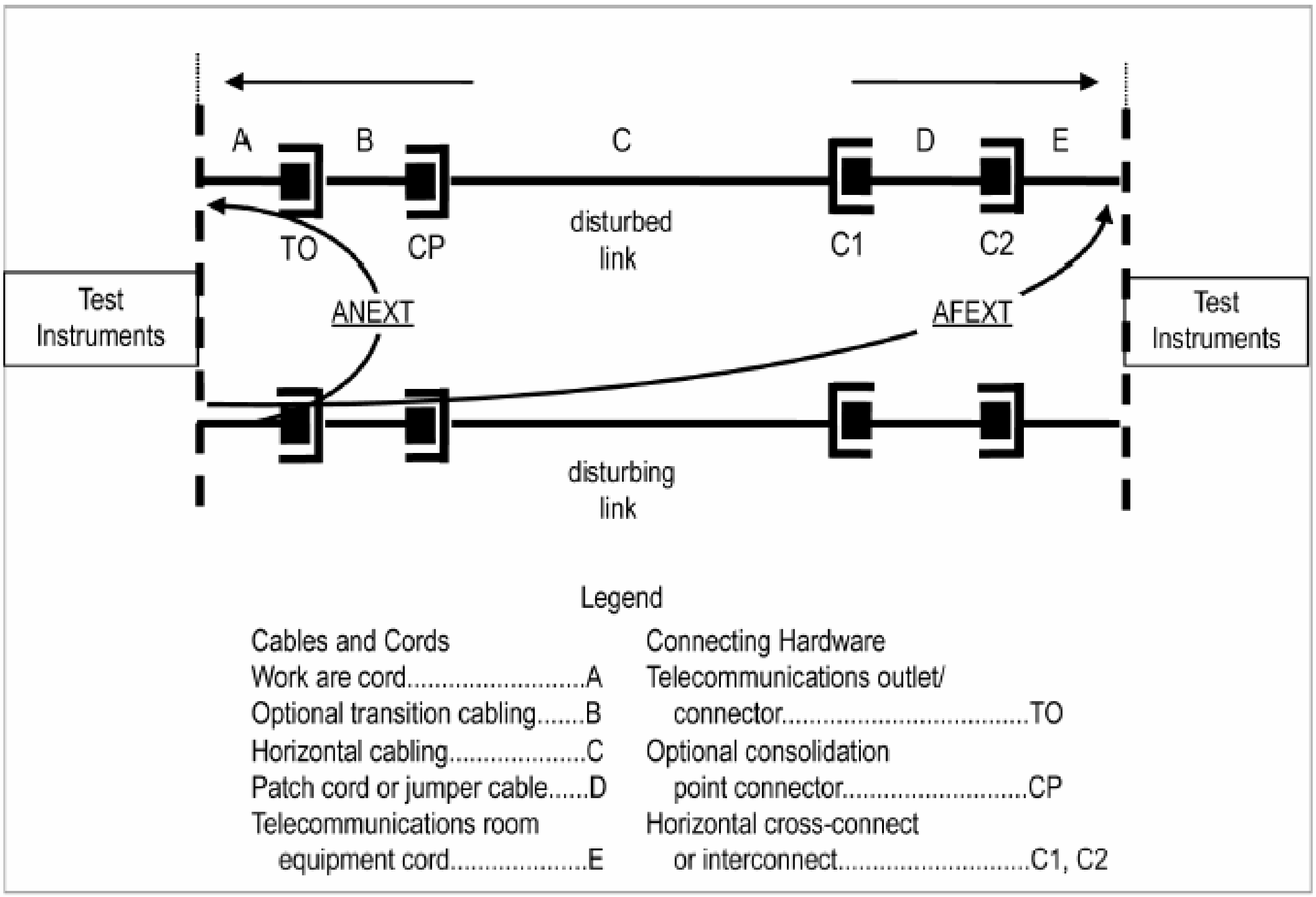
The alien crosstalk noise levels are dependent on the number and proximity of adjacent cables and connectors. The alien crosstalk is reduced as the distance between adjacent cables and connectors increases.

From an application perspective, due to the wide bandwidth of 10GBASE-T signals, the main concern is the alien crosstalk noise from other 10GBASE-T links. The relative contribution from other sources, such as 1000BASE-T, is very small. In addition, increasing the length of separated patch cords reduces the coupling voltage and reduce the coupled noise that appears back at the receiver. Possible field-mitigation strategies include:

- Selective deployment of 10GBASE-T utilizing non-adjacent patch panel positions (adjacency should also be checked at rear of the patch panel);
- Reduce the alien crosstalk coupling in the first 5 to 20 meters of the horizontal cabling by separating the equipment cords and the patch cords and un-bundling the horizontal cabling. A significant portion of the ANEXT coupling occurs in the first 20 meters of cabling;
- Utilize equipment cords sufficiently specified to mitigate the alien crosstalk coupling such as Category 6 ScTP and Augmented Category 6;
- Reconfigure the cross-connect as an interconnect; and
- Replace connectors with Augmented Category 6 connectors.

Alien Crosstalk Test Configuration

Alien crosstalk is tested between a disturbed link and one or more disturbing links in close proximity. The example test configuration for ANEXT and AFEXT shown in Figure 15 is a schematic representation of a single disturbed link and a single disturbing link.



Source: Solarflare Communications, Inc. and TSB-155

Figure 15 – Alien Crosstalk Test Configuration

Alien Crosstalk Field Measurements

Frequently, it is not practical to measure and analyze in the field alien crosstalk on all the possible cabling combinations that can be found. The overall field test strategy therefore consists of identifying worst case bundling conditions (from cabling topology through visual inspection) and then testing a sufficient number of ports to assure performance using statistical techniques.

The general strategy for field testing is as follows. First, thorough visual inspections of the topology identify the worst case bundling conditions. Then test a sufficient number of ports and use statistical techniques to assure the necessary performance.

The following list describes the method of testing and analysis when the disturber channels are within the same cable bundle as the disturbed channel; and when the channels are next to the disturbed channel where a large number of cables join (e.g. a patch panel):

- Disturber channels within the same cable bundle as the disturbed channel can contribute significant amounts of alien crosstalk. Contributions from disturbers in other cable bundles are generally not significant and need not be tested unless the cables terminate on connectors that are located in close proximity. (Note: Most of the alien NEXT coupling occurs within the first 20 m of the near-end of the cabling.);
- Start testing alien NEXT using channels next to the disturbed channel where a large number of cables join (e.g., a patch panel). Add disturber channels to the result as long as it appears that the PSANEXT result is affected. This continues until all likely disturber channels (from knowledge of the cabling topology) have been measured: and
- Start testing alien FEXT loss using disturber channels next to the disturbed channel where a large number of cables join (e.g., a patch panel). Add channels to the result, as long as it appears that the PSAFEXT loss result appears to be no longer affected. This continues until all likely disturber channels (from knowledge of the cabling topology) have been measured.

If the channel does not meet either the PSANEXT or PSAELFEXT limits independently, then the overall alien crosstalk margin computation (ACMC) is performed on the alien crosstalk measurement data. This computation looks at the combined alien crosstalk noise and compares this noise to the alien crosstalk noise budget for 10GBASE-T in the frequency range from 10 MHz to 400 MHz. The ACMC needs to be greater than zero for the worst pair and for the average of all four pairs in the disturbed channel.

Conclusion

With the advantages described in this paper, it should be expected that 10GBASE-T joins its older siblings – 10BASE-T, 100BASE-TX and 1000BASE-T – as the dominant networking technology for its speed. This paper should prove invaluable for any user, producer or student of 10GBASE-T technology.

The cost-effectiveness and plug-and-play simplicity of copper structured cabling and 10GBASE-T technology enables the rapid adoption of 10 Gigabit Ethernet into data centers and server farms, switch-to-switch connections in the core and distribution networks. Constrained by cost, the current 10 Gigabit market deployment has been generally limited to link aggregation of multiple Gigabit Ethernet downlinks and point-to-point links for long-haul high-end applications. With 10GBASE-T, network managers can now scale their networks to 10 Gigabit speeds while leveraging their investment in the installed copper cabling infrastructure. For upgrades and new installations, network managers can continue to deploy twisted pair copper cabling throughout their enterprise maintaining the benefits of seamless connectivity. 10GBASE-T promises to be a very successful high-speed networking solution for data center and enterprise networks.

Appendix A: TR42-10GBASE-T Projects

The Telecommunications Industry Associations (TIA) TR-42 Engineering Committee, which is responsible for maintaining telecommunications standards for copper and optical fiber cabling, initiated two projects in support of the development of 10GBASE-T:

- TR42 PN-3-0134 to develop a telecommunications system bulletin (TSB-155) in support of 10GBASE-T over installed cabling. TR-42 has approved the release of TSB-155. The expected publication was January 2007.
- TR42 Project designated SP-3-4426 to address new cabling specifications designated Augmented Category 6 cabling. The expected publication date is 2007.

10GBASE-T on Installed Cabling – TR42 Project ANSI/TIA/EIA-TSB155

Recognizing that 10GBASE-T is dependent on cabling performance characterized to higher frequencies than specified for Category 6 and the alien crosstalk performance of the cabling, TSB-155 recommends deploying 10GBASE-T over installed 100 Ohm , 4-pair Category 6 cabling including:

- Additional cabling guidelines for 4-pair 100 Ohm Category 6 cabling for 10GBASE-T applications;
- Extended frequency transmission performance of Category 6 cabling from 250 MHz up to 500 MHz; channels and permanent links;
- Alien near-end crosstalk (PSANEXT) and power sum attenuation to alien crosstalk ratio far-end (PSAACRF) (Note: The definition of power sum attenuation to alien crosstalk ratio far-end (PSAACRF) as used in this standard corresponds to PSAELFEXT in IEEE Std 802.3an-2006 Standard on 10GBASE-T. This standard also uses AACRF; (attenuation to alien crosstalk ratio far-end) which corresponds to AELFEXT in IEEE

802.3an-2006 Standard on 10GBASE-T.);
- Additional guidelines for field test equipment, field test methods and alien crosstalk testing;
- Field test configurations for alien crosstalk measurements of channels and permanent links;
- Alien crosstalk margin computation;
- Internal parameter mitigation; and
- Alien crosstalk mitigation procedures.

TSB-155 Distances

TSB-155 provides additional guidelines for 10GBASE-T supported distances that recognize harsh alien crosstalk environments such as the use of kellum grips in riser applications. The kellum grip provides a method of securing cable bundles in the riser space between floors as illustrated in figure 14. The cables are bound tightly by the grip. The kellum grip can increase the alien crosstalk coupling compared to other methods of binding the cables such as tie-wraps.

For cabling configurations in harsh alien crosstalk environments such as those utilizing kellum grips, TSB-155 states 10GBASE-T operational channel lengths of up to 37 meters over Category 6 cabling.

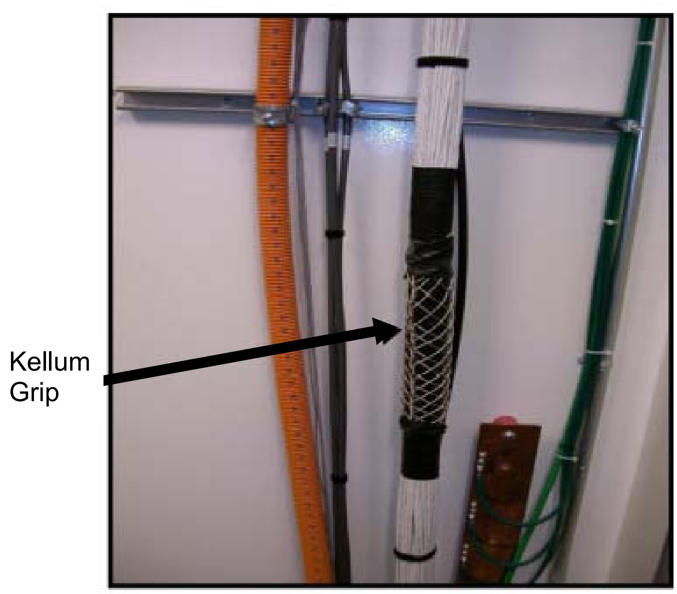


Figure 16 – Kellum Grip, Vertical Cable Bundle Support

10GBASE-T on New Cabling: TR42 Project: ANSI/TIA/EIA-568-B.2-10

Project SP-3-4426-AD10 was initiated to develop cabling and component specifications and test procedures to support the operation of 10GBASE-T over 100 meters of structured balanced twisted-pair copper cabling. This effort resulted in the development of ANSI/TIA/EIA-568-B.2-10, the Transmission Performance Specifications for 4-pair Ω 100 Augmented Category 6 Cabling. These specifications cover many of the same topics as the TSB-155 document outlined above including:

- Transmission performance specifications for 4-pair Ω 100 augmented Category 6 cabling (Category 6A);
- Cabling specifications, (cables and connecting hardware);
- Internal cabling parameters (insertion loss, return loss, near-end crosstalk (NEXT) loss, equal level far-end crosstalk (ELFEXT), propagation delay, delay skew, worst pair-to-pair and power sum);
- Accuracy requirements for level IIIe field testers;
- Alien near-end crosstalk (PSANEXT) and power sum attenuation to alien crosstalk ratio far-end (PSAACRF) (Note: The definition of power sum attenuation to alien crosstalk ratio far-end (PSAACRF) as used in this standard corresponds to PSAELFEXT in IEEE Std 802.3an-2006 10GBASE-T. This standard also uses AACRF (attenuation to alien crosstalk ratio far-end) which corresponds to AELFEXT in IEEE Std 802.3an-2006 10GBASE-T.);
- Alien crosstalk mitigation not required to support 10GBASE-T;
- Near-end crosstalk (NEXT) loss and return loss requirements for modular plug cords; and
- Balance requirements for connecting hardware and cables.

Appendix B: An Overview of IEEE Std 802.3an-2006 Standard

Background

In June 2002, the first 10 Gigabit Ethernet standard, IEEE Std 802.3ae-2002, was ratified by the IEEE Standards Association (IEEE-SA) Standards Board. It heralded the ability of Ethernet to interoperate at speeds of 10 Gb/s; however, the focus was entirely optical. Ethernet technology has had a long tradition of operating over copper cabling, especially unshielded twisted-pair (UTP) copper cabling. Discussions about how to transport 10 Gigabit Ethernet over UTP copper cabling began in the late summer of 2002.

In November 2002, an IEEE 802.3 call-for-interest (CFI) occurred to request the formation of a 10 Gigabit Ethernet over UTP copper cabling study group, dubbed 10GBASE-T. With strong support of the IEEE 802.3 working group (WG), the study group was formed with the charter to draft a set of objectives, respond to the five criteria and develop a project authorization request (PAR).

Project Scope

The scope of the project was pretty simple: transmission of 10 Gigabit Ethernet over unshielded twisted-pair cabling. Together with two of the primary cabling standards bodies, TIA TR-42 and ISO/IEC JTC 1/SC 25/WG 3, it became apparent that there was no support for modifying Category 5, Category 5e or Class D cabling specifications. The study group therefore decided to focus its efforts on better cabling. Category 6 and Class E cabling became the minimum requirement for 10GBASE-T operation. Other cabling types are also supported so long as they meet the channel requirements.

Objectives

With the minimum set of cabling specifications required, the study group was able to set the following objectives for the project:

- Preserve the 802.3/Ethernet frame format at the MAC Client service interface;
- Preserve minimum and maximum frame size of current 802.3 Std;
- Support full duplex operation only;
- Support star-wired local area networks using point-to-point links and structured cabling topologies;
- Support a speed of 10.000 Gb/s at the MAC/PLS service interface;
- Select copper media from ISO/IEC 11801:2002, with any appropriate augmentation to be developed through work of 802.3 in conjunction with SC25/WG3;
- Support Clause 28 auto-negotiation;
- Support coexistence with 802.3af;
- Do not support 802.3ah (EFM) OAM unidirectional operation;
- Meet CISPR/FCC Class A;
- Support operation over 4-connector structured 4-pair, twisted-pair copper cabling for all supported distances and classes;
- Define a single 10 Gb/s PHY that would support links of:
- At least 100 m on four-pair Class F balanced copper cabling;
- At least 55 m to 100 m on four-pair Class E balanced copper cabling; and
- Support a BER of 10^-12 on all supported distances and Classes.

The first five objectives were specified to ensure that the 10GBASE-T physical layer device (PHY) would support the existing 10 Gigabit Ethernet MAC, XGMII and XAUI that were specified in IEEE Std 802.3ae-2002. Selecting copper media from ISO/IEC 11801 gave the study group a base reference from which to start. The study group also knew in advance that the ISO/IEC JTC 1/SC 25/WG 3 (along with TIA TR42) would be working on extending the specification of Class E cabling to support 10GBASE-T operation, plus there was the possibility that they would better cabling specifications.

Clause 28 auto-negotiation is the same that is used in 10/100/1000 Mb/s Ethernet, and support of that protocol was considered a requirement. With the ratification of IEEE Std 802.3af-2003 data terminal equipment (DTE) power via the Ethernet cable, the study group decided to support coexistence with IEEE Std 802.3af-2003. An IT professional may have 10GBASE-T and DTE power in the same network or data center therefore the ability to coexist was extremely important.

During the development of the standard the committee conducted an extensive analysis of the media capabilities and the noise environments. This analysis enables users to understand the limitations as well as the capabilities of 10GBASE-T. In addition to running on Class F cabling (Category 7), and Class E cabling (Category 6), cabling experts set about specifying a new category of cable optimized for 10GBASE-T; augmented Category 6 abbreviated Category 6A specified in TIA/EIA-568-B.2-10 and Class EA specified in ISO/IEC 11801 Edition 2.1.

The IEEE Std 802.3ah-2004 Ethernet in the First Mile Operations Administration and Maintenance (OAM) unidirectional mode was to support management features in access networks and was deemed unnecessary for local area networks. The CISPR/FCC requirements and BER requirements are basic requirements for operation. The CISPR/FCC requirements relate to electromagnetic interference (EMI) that equipment should be able to comply with. The BER of 10-12 was selected to be consistent with optical 10 Gb/s Ethernet specified in IEEE Std. 802.3ae-2002.

The remaining objectives specified the cabling that would be considered as the basis for the standard. A range was documented for Class E because depending on the installation practices or the cable construction the reach capabilities could change. For example, tightly bundled Class E (or CAT6) cabling would be susceptible to more crosstalk from the other cables in the bundle, thereby shortening the achievable reach.

Requirements

The study group was also responsible for generating a five-criteria document that explained the following basic requirements:

- Broad market potential;
- Compatibility with IEEE standard 802.3;
- Distinct identity;
- Technical feasibility; and
- Economic feasibility.

Broad Market Potential

Broad market potential needs to show a broad set of applications, support by multiple vendors, interest from multiple users, and a balanced cost between the network and the attached stations. The standards effort planned to focus on the data center where data aggregation was demanding higher bandwidth, and with the cost and infrastructure support offered by a 10 Gigabit Ethernet twisted-pair copper cabling solution. There was strong vendor and user support with almost 70 people from over 30 companies expressing an interest in participating.

Compatibility with the IEEE 802.3 Standard

Compatibility with 802.3 and distinct identity were pretty easy as the project planned only to focus on the physical (PHY) layer and use the XGMII or XAUI to interface to the already specified 10 Gigabit Ethernet MAC. No changes were going to be from XGMII/XAUI on up the layers.

Distinct Identity

As for being distinct, this was the only PHY that was to be specified for 10 Gigabit Ethernet transmission on four-pair, twisted-pair balanced copper cabling.

Technical Feasibility

Proving technical feasibility was more of a challenge. It was very difficult to demonstrate feasibility of a technology 3-4 years in advance. Ethernet in the past has had the opportunity to borrow existing technologies e.g., FDDI for Fast Ethernet, and Fibre Channel for Gigabit Ethernet. However, because the challenge of increasing the bit rate in a channel is universal, 10GBASE-T was able to leverage existing technologies.

Economic Feasibility

The final criterion was economic feasibility. Knowing cost factors and predicting the cost of performance is always speculative. The group agreed that a 10 times performance increase was possible with only a three- to four-times increase in cost because of savings associated with cabling and installation.

PAR Approval

In addition to the objectives and the five criteria, the group prepared the project authorization request (PAR). The PAR is a request to form a project to generate a standard. The PAR describes the name of the project, which group is doing it, the contacts and the purpose and scope of the project. Once completed all these documents were submitted to 802.3 with only the PAR and 5 Criteria going to the 802 executive committee, and only the PAR going to the IEEE-SA new standards committee (NesCom) who review and recommend to the IEEE-SA Standards Board if it should be approved or not.

With the PAR approved by 802.3, the 802 executive committee (EC) and the IEEE-SA standards board, the study group became a task force to develop a draft standard for 10GBASE-T. The project received the designation IEEE P802.3an.

Drafts and Approval

Once the IEEE P802.3an Task Force was approved, the first task was to adopt baseline proposals from which a first draft could be created. Those proposals including things such as the modifications required to Clause 28 auto-negotiation to the use of low density parity check (LDPC) framing and double-square 128 (DSQ128) constellation mapping. Once the basic framework of the draft standard was created, draft D1.0 was issued for task force review. During this review process, the draft standard progresses to a point of becoming technically complete. In other words, the technical framework for a 10GBASE-T PHY is incorporated into the draft.

Once completed, draft D2.0 was created and released to the 802.3 WG for balloting and comments. Balloting is used to ensure that there is 75% approval of the WG respondents. In the 802.3 WG ballot phase, the goal of the task force is to make sure that the draft standard is technically correct. Where the task force review creates the framework, the WG ballot fills in the details. When completed, all that remains is to "buff and polish" the draft during sponsor ballot.

At the end of sponsor ballot, the task force requested approval to submit the draft standard to the IEEE-SA review committee (RevCom) for ratification. If the approval is received from the 802.3 WG and the 802 EC, then the IEEE P802.3an draft standard is placed on the RevCom agenda. On June 7, 2006, RevCom recommended the approval of the draft. On June 8, 2006, the IEEE-SA standards board approved the project and the draft was ratified and renamed to be IEEE Std 802.3an-2006 Standard.

Glossary

alien crosstalk Unwanted signal coupling from a disturbing pair of a 4-pair channel, permanent link, or component to a disturbed pair of another 4-pair channel, permanent link, or component.

alien far-end crosstalk (AFEXT) The unwanted signal coupling from a disturbing pair of a 4-pair channel, permanent link, or component to a disturbed pair of another 4-pair channel, permanent link, or component, measured at the far-end.

alien near-end crosstalk (ANEXT) Unwanted signal coupling from a disturbing pair of a 4-pair channel, permanent link, or component to a disturbed pair of another 4-pair channel, permanent link, or component, measured at the near-end.

attenuation to alien crosstalk ratio at the far end (AACRF) The difference in dB between the alien far-end crosstalk from a disturbing pair of a 4-pair channel, permanent link, or component and the insertion loss of a disturbed pair in another 4-pair channel, permanent link, or component.

average power sum alien near end crosstalk loss The calculated average of the power sum alien near-end crosstalk loss of the four pairs of the disturbed channel or permanent link.

average power sum attenuation to alien crosstalk ratio far-end The calculated average of the power sum attenuation to alien crosstalk ratio far-end of the four pairs of the disturbed channel or permanent link.

power sum alien far-end crosstalk (PSAFEXT) The power sum of the unwanted signal coupling from multiple disturbing pairs of one or more 4-pair channels, permanent links, or components to a disturbed pair of another 4-pair channel, permanent link, or component, measured at the far-end.

power sum alien near-end crosstalk (PSANEXT) The power sum of the unwanted signal coupling from multiple disturbing pairs of one or more 4-pair channels, permanent links, or components to a disturbed pair of another 4-pair channel, permanent link, or component, measured at the near-end.

power sum attenuation to alien crosstalk ratio at the far end (PSAACRF) The difference in dB between the power sum alien far-end crosstalk from multiple disturbing pairs of one or more 4-pair channels, permanent links, or components, and the insertion loss of a disturbed pair in another 4-pair channel, permanent link, or component.

Note: The definition of power sum attenuation to alien crosstalk ratio far-end (PSAACRF) corresponds to PSAELFEXT in IEEE Std 802.3an-2006 10GBASE-T. In addition, AACRF (attenuation to alien crosstalk ratio far-end) corresponds to AELFEXT in IEEE Std 802.3an-2006 10GBASE-T.