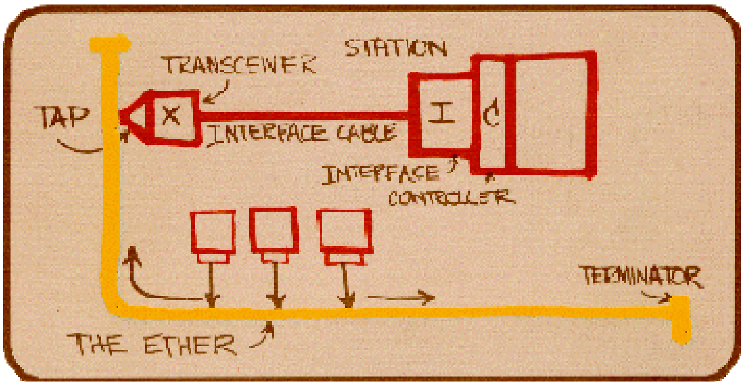
10G Ethernet IEEE 802.3an

By Allan Nielsen

1. Introduction

Ethernet continues its successful dominance within the communication industry, starting back in 1976 as a concept created by Robert M. Metcalfe and presented to the National Computer Conference in June 1976.



From its first edition targeting 1 Mbps Ethernet has taken the lead in the communication industry, as speeds are soon to hit 100 Gbps using thin 65nm technology chipsets.

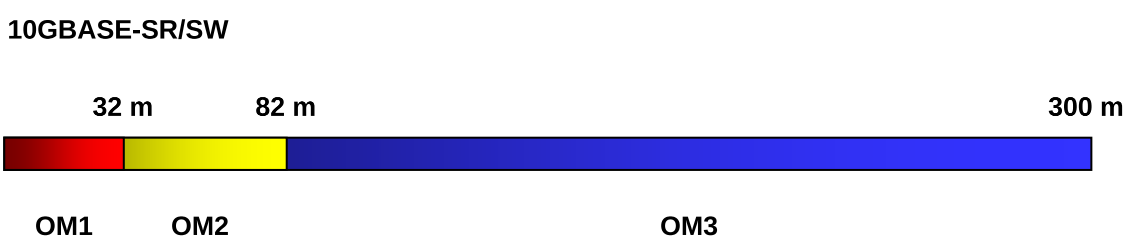
This document will discuss the lifetime and progress of Ethernet together with the suitable media needed for its transmission.

2. Ethernet

Ethernet in the 3rd millennium has taken quantum leaps from state of the art in the 1990's which was 1000 BASE (1 Gbps) with the introduction of 10G BASE (10 Gbps) in 2002.

Originally 10G BASE was designed for fibre optic backbone systems, using low dispersion Multi Mode fibre, known as OM3 fibre or Singlemode fibre. Reaching an age where Ethernet was limited by the available bandwidth in the cabling, rather thanby the length and attenuation of the cabling, the application started a new era in communications.

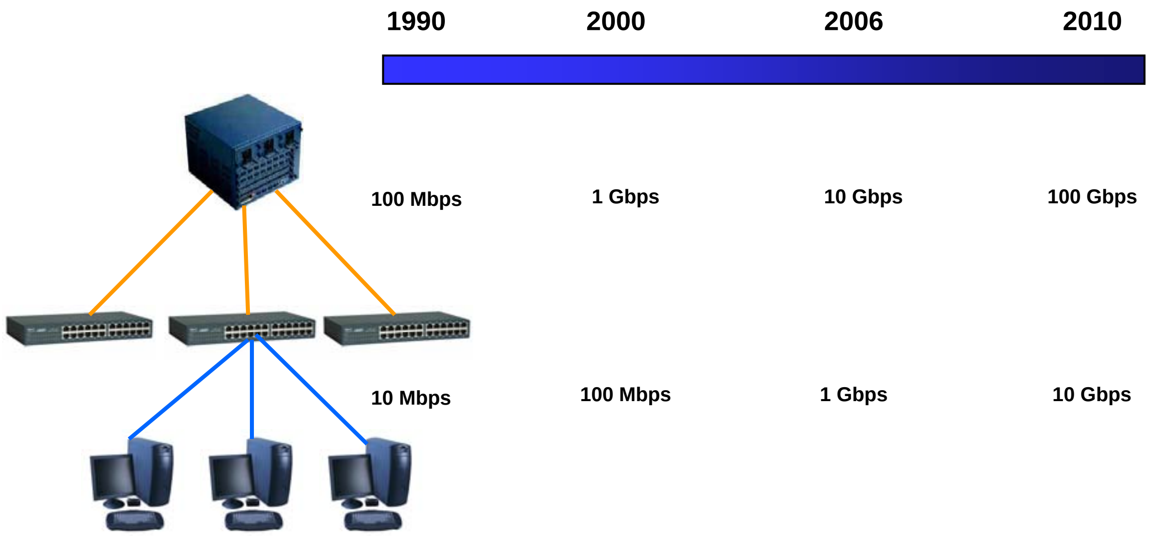
Bandwidth limited fibre which has been installed throughout the world was suddenly de-rated to a second class fibre optical mediim with the introduction of 10G BASE over fibre; even at 1000 BASE SX the bandwidth of OM1 fibre reduced the supported length to 275 meters. 10G Ethernet reduced the supported length even further:



Only bandwidth optimized fibre such as OM3 supports applications like 10GBASE-SR/SW up to a usable length, leaving older fibres as OM1 and OM2 almost obsolete.

3. Speed

Most new computers today are equipped with a 1000BASE-T interface on the motherboard, which means that most of the workspaces throughout office areas are Gigabit ready. Problems are created in the aggregation speed used in backbones, as these often are not upgraded at the same speed as the work area:



It is common sense to use a higher speed as the aggregation speed, meaning that backbone systems always have to be one level higher than the horizontal work area.

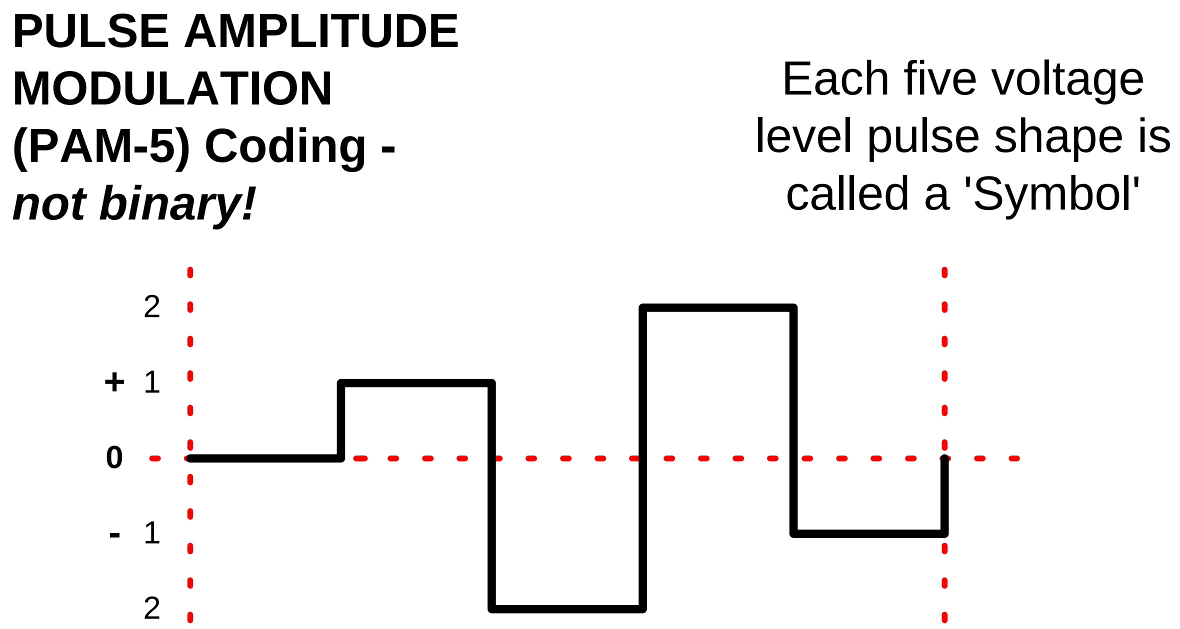
As 100 Mbps Ethernet is the lowest common speed in the horizontal work area, we are approaching an age where 1 Gbps will be the lowest common speed, and for that we need to consider backbone systems with speeds of 10 Gbps or even higher.

4. Coding structure

In recent years the communication industry has developed methods of using 10G BASE with twisted pair media; this has created a whole new set of requirements for balanced cabling as we know it today.

The new application uses a high density coding structure which enables several bits to be transmitted in one frequency cycle; already with the introduction of 1000BASE-T a complex coding structure was used:

PAM 5:



Logically even at 1 Gigabit Ethernet the noise sensitivity increased significantly from previous speeds, as the coding of one frequency cycle contained more and more information.

With the emerging higher speed of 10 Gigabit Ethernet on copper, a new coding structure had to be implemented, as a PAM 5 coding would not provide the required amount of transmission data on the frequency provided by cabling systems. Therefore IEEE developed a new coding structure DSQ 128:

10GBASE-T is defined by IEEE 802.3an. It uses 4-pair transmission (shown in figure 1) and is designed to support up to 100 metres of Category 6_A or higher cabling. The aggregate data rate of 10 Gbit/s is achieved by transmitting 2.5 Gbit/s over each cable pair. Hybrids and cancellers are used to facilitate simultaneous transmission in both directions (i.e. full duplex).

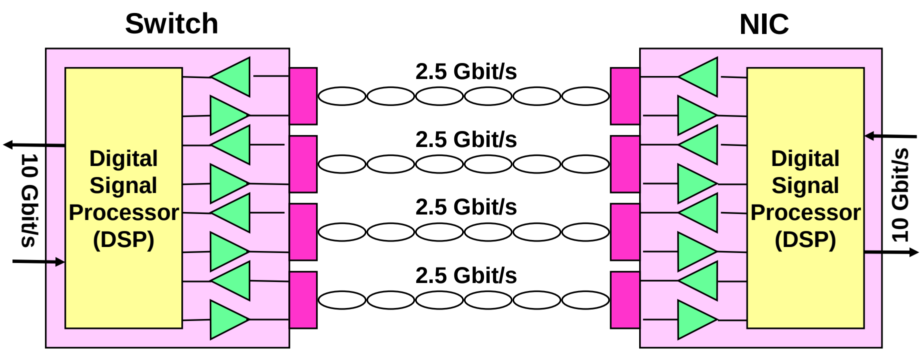


Figure 1: 4-pair full-duplex transmission used by 10GBASE-T

10GBASE-T uses 16-level PAM signalling in each pair. 16 levels of Pulse Amplitude Modulation are derived from 4 binary bits ($\log_2 16$), and each 16-level PAM value is referred to as a PAM16 symbol. A two dimensional (2D) code is created using a pair of PAM16 symbols. This produces a constellation of 256 (16x16) orthogonal values. Double SQuare 128 (DSQ128) is then constructed by pruning alternate points from the 16x16 array to produce 128 2D symbols (this is like a checkerboard with black & white squares). The resulting DSQ128 constellation is illustrated in figure 2 (this uses values -15, -13, -11, -9, -7, -5, -3, -1, +1, +3, +5, +7, +9, +11, +13, +15 in two dimensions).

Separation between adjacent DSQ128 code values is increased by √2 compared with a conventional 16x16 orthogonal array; this provides 3dB improvement in signal-to-noise ratio and leads to a significant reduction in bit-error-rate (BER).

The modulation rate of PAM16 symbols is 800 MSymbols/second (800Mbaud), which puts the Nyquist frequency for baseband signalling at 400 MHz. PAM codes generally require 10-15% headroom above the Nyquist frequency for practical implementations, and 10GBASE-T cabling is specified with an upper frequency of 500 MHz to provide an adequate margin for transmission.

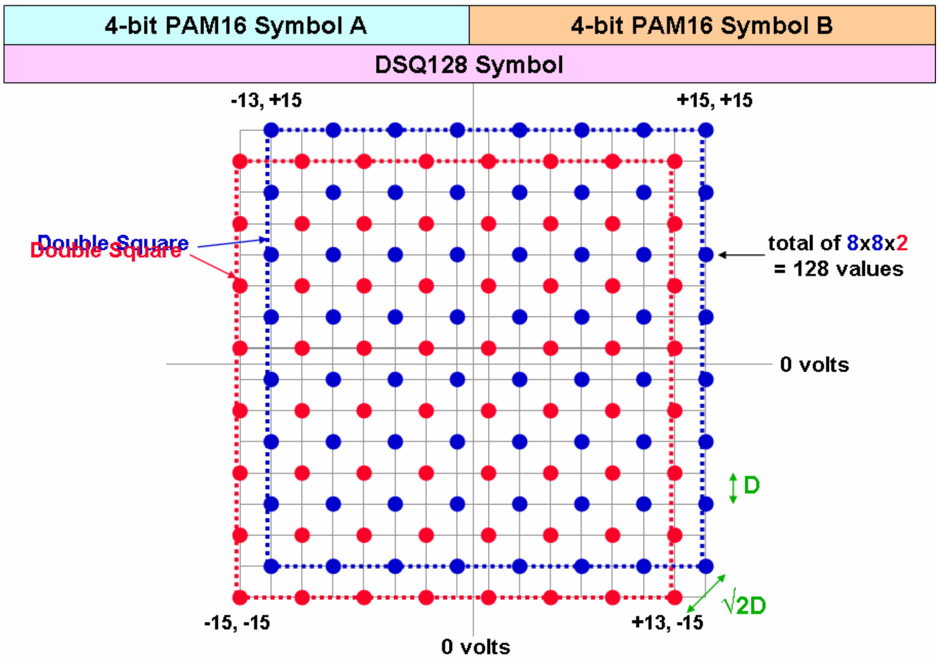


Figure 2: DSQ128 Constellation

The constrained DSQ128 constellation is determined by 7-bit labels ($\log_2 128$), each comprising 4 coded data bits plus 3 uncoded data bits. Coded data bits are protected by a Low Density Parity Check (LDPC) code. Code blocks of 512 DSQ128 labels are processed by the 10GBASE-T Physical Layer. Bit mapping of these code blocks is shown in figure 3. 325 check bits are used to protect the integrity of each block of 1723 coded data bits to form an LDPC codeword of 2048 bits. This is one of many LDPC codes and is referred to as LDPC (1723, 2048).

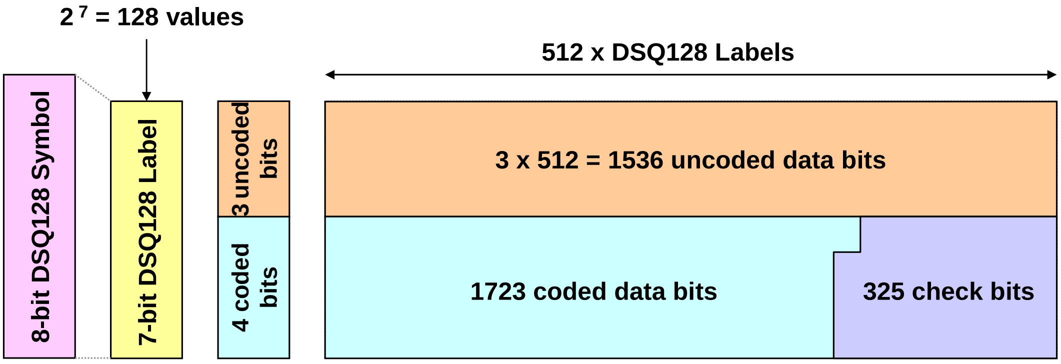


Figure 3: DSQ128 Code block with LDPC (1723, 2048)

Framing and code block overheads need to be taken into account when calculating the data throughput of a 10GBASE-T system. If we first examine the information payloads contained in each DSQ128 symbol;

- 4 bits of coded data; however the average number of data bits per DSQ128 symbol is reduced by the 325 check bits use in each 2048 bit LDPC code block. This is reduced by 1723 ÷ 2048 to give 3.3652 bits.
- 3 bits of uncoded data with no overhead.
- The average number of data bits per DSQ128 symbol is therefore 3 + 3.3652 = 6.3652 bits.
- As two PAM16 symbols make a DSQ128 symbol, the average number of data bits per PAM16 symbol is half of this, or 3.1826 bits.

Additional overheads exist at the Physical Layer for 64B/65B block coding, frame synchronisation and Cyclic Redundancy Checksums (CRCs), thus reducing the average number of bits per PAM16 symbol to 3.125 bits.

The data throughput for 10GBASE-T is then calculated as;

800 MSymbols/second x 3.125 bits/symbol x 4 pairs = 10 Gbits/second

Although complex, the choice of DSQ128 makes a significant contribution to the throughput and robustness of 10GBASE-T, so helping to achieve its transmission quality objective with a bit error rate of no worse than 10^{-12}.

5. Impact on cabling systems

The DSQ128 coding structure requires the bandwidth of the cabling system to be approximately 417MHz, which increases the requirements for the cabling system beyond the requirements given for Class E / Category 6 cabling. As the application group of IEEE wanted to support both a UTP and STP cabling infrastructure, they had to compensate for quite some noise created in the cabling system.

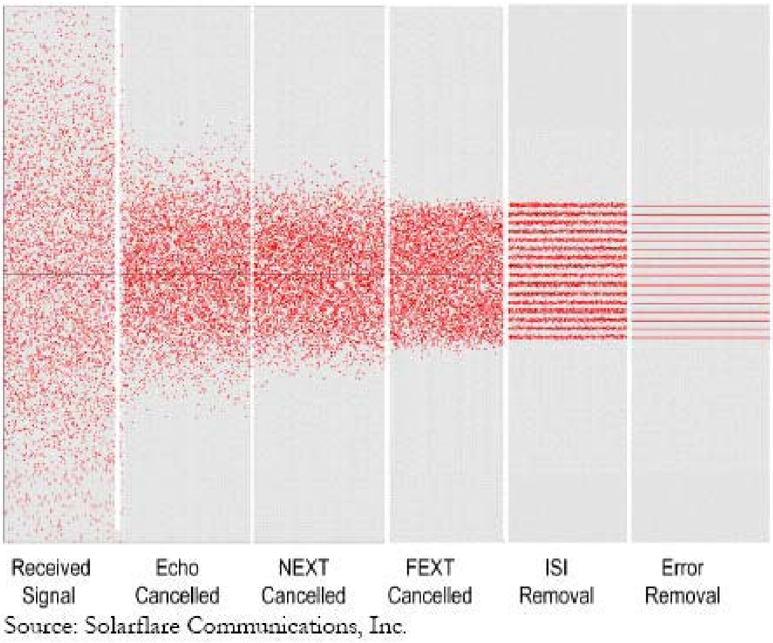
The cancellation of Near End Crosstalk (NEXT) was required to an extent not yet seen before, as this parameter has a negative signal to noise ratio, also know as Attenuation to Crosstalk Ratio (ACR) in the cabling standards. The requested bandwidth for 10GBASE-T exceeds the positive signal to noise ratio for Category 6 cabling by 200 MHz and therefore cancellation techniques have to be built into the application.

NEXT can be cancelled within the same transmission channel, as the transmitter knows what signals it has sent, and therefore knows that echoes of this signal can be removed. However this requires some processing power of the PHY, as it always has to monitor its incoming signals for similar signals to those just transmitted.

The cancellation of Far End Crosstalk (FEXT) was also implemented with the use of a similar technology for the far end. As the far end receiver does not know the signal to be received, it must rely on the signals received, and then sort out what is the real signal and what is the noise signal. This is done by monitoring on the strongest signal received, as this often arrives some milliseconds before the noise signal, and therefore a noise cancellation processing function is able to sort out the real signal.

Reduction of Echo is also needed. This is perhaps better know as Return Loss in the cabling business, which relates to the echoes received on the same transmitting pair when transmitting and receiving at the same time. Similar to the cancellation of NEXT, this is done by processing power, as the transmitter knows what signals that have just been transmitted, and is therefore able to cancel the echo noise.

Finally an intersymbol interference correction is done by the digital processing of the received signal, which removes any further errors from the transmitted signal.

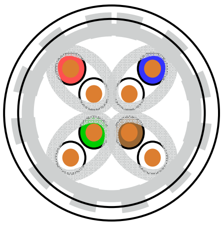


Source: Solarflare Communications, Inc.

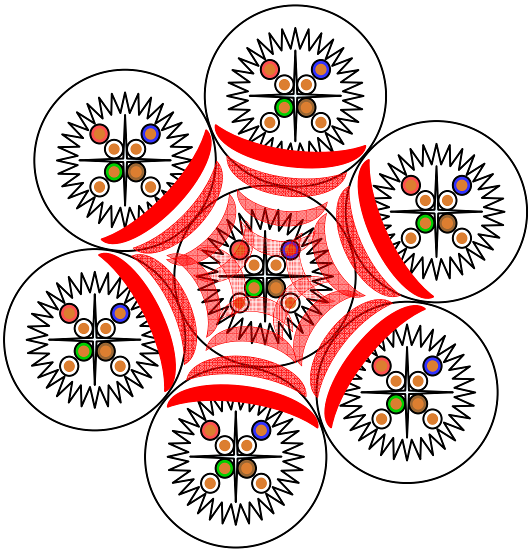
The things that the digital signaling processing cannot remove are noise from unknown sources, such as adjacent signalling ports or background noise. Therefore it was necessary to specify the minimum performance for these parameters, which was done by the addition of Alien Crosstalk and background noise assumptions.

Alien Crosstalk, which consists of both Power Sum Alien Near End Crosstalk and Power Sum Alien Far End Crosstalk, is defined as the noise coupled from adjacent transmission channels / cables into the used transmission channel, and as these signals are unknown (Alien) or Exogenous they cannot be compensated; therefore the new generation of cabling infrastructure needs to compensate for these unwanted noise couplings.

The cancellation of Alien Crosstalk can be achieved in general bytwo technologies; the first and simplest is shielding, which has been used around the world to reduce background noise and unwanted signals in transmission channels for decadesand is commonly known by engineers who design critical transmission lines in for example cars and airplanes, where a disturbance in the signal could have severe consequences. The shielding technology can be divided into two major groups: the first one is a simple F/UTP cable where the 4 pairs of a transmission channel are surrounded by a foil screen which is wrapped around the 4 pairs throughout the whole transmission channel, the other technology is called "Pairs in Metal Foil" (PiMF) where each pair in the cable is covered by a foil and the 4 pairs are covered by either a braid screen or a foil screen. These are often referred to as Category 7 cables, but the technology arrived already in the early 1980's with IBM Cabling Type 1, where two pairs with their own foil screen were encapsulated by a braid screen. This technology was suppressed by unscreened cables in the late 1980's as they offered smaller cables, less restrictions and easier installation. Combined with the lower cost of the materials, this completely changed the cabling infrastructure market into the one we know today.

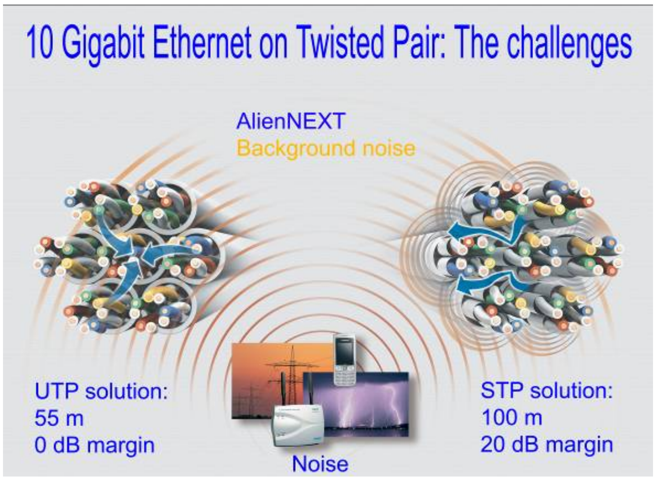


The unscreened cables for the next generation of cabling also needs to cancel out the Alien Crosstalk from adjacent transmission channels, which is achieved by a combination of pair twisting and separation of the transmission channels. Typical pair twisting will provide a 40 dB cancellation of unwanted signals, and some of the best cables are able to provide up to 60 dB cancellation of unwanted signals; but this is not enough, as the requirements for Alien Crosstalk are up to 90 dB, or 1024 times more protection than provided by the best pair twisting. The only parameter left is then separation of the cables throughout the whole installation, where a separation of up to 5 mm is needed. This can be achieved by a thicker outer jacket which is now specified up to 9 mm in diameter by the standards, which is 20% thicker than a PiMF cable. The combination of a good and very tight pair twisting and the much thicker outer jacket provides the UTP cable with the required minimum cancellation of Alien Crosstalk, but it does not provide any of the margins as provided by a shielded cabling system.



The second parameter which cannot be removed by digital signal processing is background noise, which means signals being transmitted from wireless services such as television signals, wireless Ethernet or cellular phones.

An unscreened cabling system suffers in regard to background noise: whereas a screened cabling system is naturally protected against background noise from the screening effect of the foil and braids, an unscreened system needs to have additional protection from aerial transmitted signals such as television, WiFi and cordless phone, which all transmit in the same frequency specter as used by 10GBASE-T. This can only be achieved by a fully screened pathway, expanding from the wiring closet all the way to the wall outlet.



With the introduction of background noise as one of the key design parameters within cabling systems, EMC becomes of more importance than before. The EMC performance is measured as Coupling Attenuation, which is the cabling system's ability to attenuate a noise coupling from an adjacent noise source. UTP systems must have a minimum 40 dB Coupling Attenuation, while the state of the art systems provide approximately 50 to 60 dB and a good PiMF system provides well above 120 dB of protection.

With the minimum requirement for background noise of 10GBASE-T, which is -150 dB, the challenges really start for UTP cabling systems, as even with the state of the art cabling systems it cannot provide enough protection against unwanted signals such as aerial transmitted television signals, which inside buildings often can be found at the level of -80 dB or higher. This requires a UTP cabling system to be protected all the way from the wiring closet to the wall outlet, in order to reach the minimum required background noise level as described in the application standard.

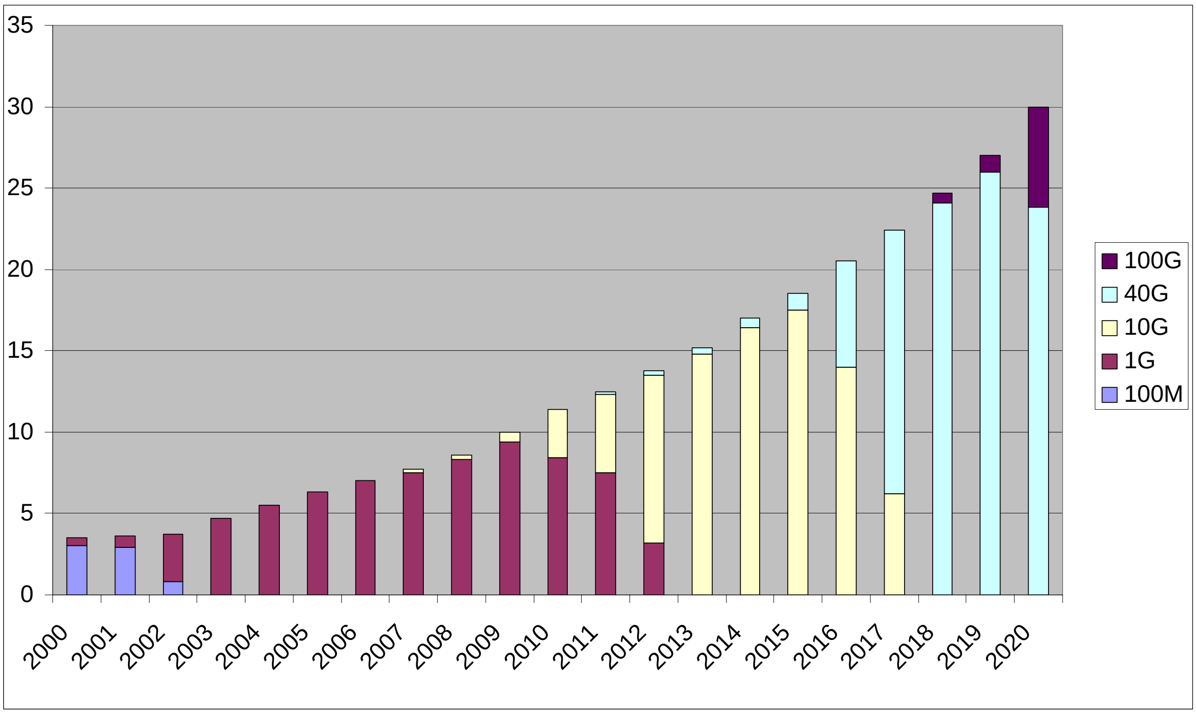
Regardless of these hurdles, it is foreseen that 10GBASE-T will dominate the LAN market for the next 10 years, and the successor applications such as 40G and 100G Ethernet will start to replace 10GBASE-T.

6. Market size

Up to the end of 2007 there was no release of a fully compliant switch or NIC card for 10GBASE-T, although market research done by IDC shows that 10G is already deployed in Data Centre environments, but these are deployed using 10G optical and CX4 twinaxial ports.

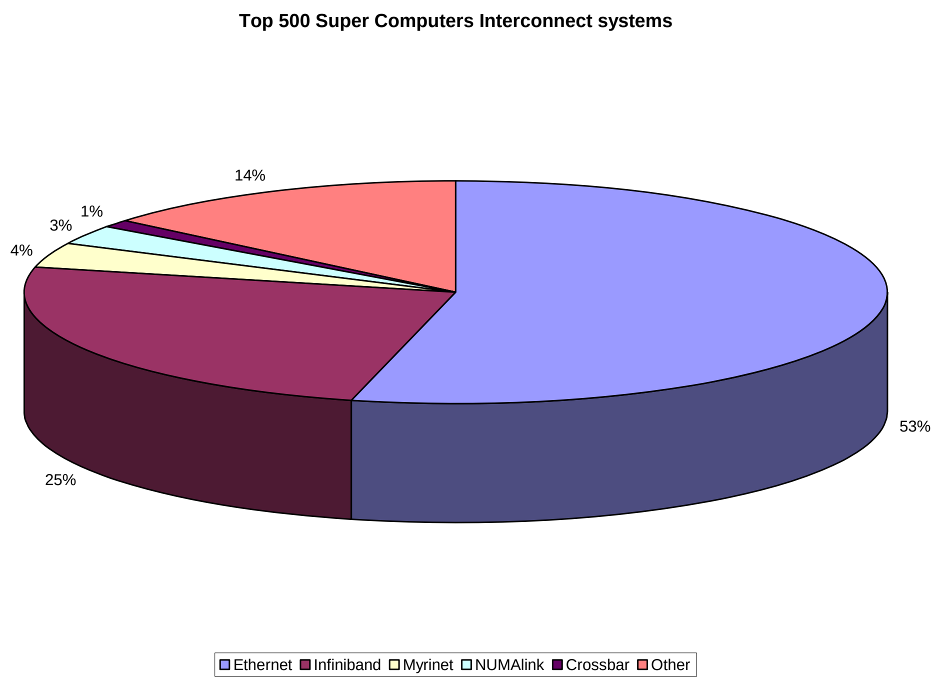
From 2001 to 2007 a relatively small number of 10G Ethernet ports were shipped from manufacturers of 10G interfaces, accounting for less than 350.000 ports, consisting of 10G optical and 10G CX4 ports. The expectation from the introduction of 10G over twisted pair is quite impressive compared to what has already been shipped to date.

With the introduction of 10GBASE-T Network Interface Cards in early 2008, it is expected that 10G Ethernet will dominate the server industry within the next 3-5 years, where market analysis from IDC shows that more than 99% of a little more than 10 million server cards shipped will be equipped with 10GBASE-T connections.



Therefore even today itiss important that users who are planning to install cabling systems, should consider cabling systems which are capable of transmitting 10GBASE-T, taking into consideration the complexity of the signals and the new needs for noise cancellation.

Even in a data centre where many connections are made with different applications such as Infiniband and Myrinet, Ethernet now accounts for more than 53% of all connections. This is expected to grow even further as 10GBASE-T will be able to replace some of the functions that currently are running on Infiniband and Fibre Channel, such as Storage Area Networks, Supercomputer Grids etc.



Top 500 Super Computers Interconnect systems
14%
1%
3%
4%
53%
25%
Ethernet
Infiniband
Myrinet
NUMAlink
Crossbar
Other

7. Cabling standards response

Cabling standards such as ISO/IEC 11801, EN 50173 and TIA/EIA 568 have responded to these new requirements in quite different ways. First with any development in this area was TIA/EIA 568B, where the committee in charge of that standard decided to develop a new Category of cabling to support the new requirements from IEEE 802.3an.

TIA/EIA 568B.2 Amendment 10

As TIA/EIA TR42.7, which is the national committee responsible for the cabling standard TIA/EIA 568B, started to develop a new Category of cabling, they decided to create a standard which would allow UTP to meet the requirements of the cabling standard. This committee was the most active within IEEE, who lobbied to lower the requirements of the application standard to a level where UTP cabling systems would be able to meet the requirements.

Their first action point was to lower the expectations for NEXT performance for cabling systems, as it was expected that most UTP cabling systems would not be able to meet the new requirements at 500 MHz. Therefore they decided to make a breaking point at the Category 6 (250 MHz) upper frequency, from where the requirements would be relaxed from 251 to 500 MHz to accommodate for the lower performance of UTP systems.

The creation of Category 6A, which is specified up to 500 MHz, only meets the application requirement and does not provide any margins over the minimum requirements finally stated in IEEE 802.3an 10GBASE-T.

Their second action was to lower the expectations for Alien Crosstalk, as it showed that UTP cabling systems suffered with both longer lengths and shorter lengths, as these represent the worst case scenario for cabling systems. Through a series of complex arguments, maybe that's why its called "Augmented" TIA/EIA, they succeeded in convincing IEEE to lower their expectations for Alien Crosstalk by more than 25%, which created even more concerns within the application industry, as they now had to compensate even more on the NEXT, FEXT and Return Loss parameters, to compensate for the higher noise picture from the outside sources.

The whole process ended up with the creation of Amendment 10 to TIA/EIA 568B.2 – Augmented Category 6, or in short Category 6A, which was released in February 2008.

This new amendment contains the requirements for Channels, Permanent Links and components, and is considered to be the most complete standard in the market today.

ISO/IEC 11801 Amendment 1 and Amendment 2

ISO/IEC JTC1 SC25 WG3, which is the international standards committee responsible for ISO/IEC 11801, was bolder than TIA/EIA, as they extrapolated the requirements from Class E from 250 MHz to 500 MHz. This allows for the support of 10GBASE-T and provides margin to the application to ensure that it will run trouble free.

The new ISO/IEC 11801 Amendment 1 contains the channel requirements for Class EA and Class FA, both of which support 10GBASE-T with margins. This standard was influenced by both UTP manufacturers and STP manufacturers of cabling systems and is therefore more balanced in regards to the expected performance of the cabling systems.

The UTP manufacturers, who mainly comes out of the representation from TIA/EIA, tried to lower the requirements to the level of TIA/EIA, and the STP manufactures tried to increase the requirements to the level where STP could show its true performance. The final outcome was a balance between the two technologies, but the battle continues in the market place.

The new Class EA, which had to accommodate both UTP and STP solutions, was extrapolated to give the internal parameters such as NEXT, FEXT and Return Loss, while the outside influences to the cabling system, Alien Crosstalk, copied what was developed in TIA/EIA, as this would otherwise eliminate UTP from the scope of the standard.

The new Class FA on the other hand is purely specified from STP requirements, and shows the true performance of shielded systems, as all the internal parameters such as NEXT, FEXT and Return Loss greatly exceed the requirements for Class EA. This also allows this new class to support a frequency which is the double of Class EA, 1000 MHz. This allows the cabling system to support new applications such as CATV.

Amendment 1 to ISO/IEC 11801 was released in February 2008

ISO/IEC 11801 Amendment 2 will contain the missing parts of Class EA and Class FA to complete the standard as a full document, as this is going to contain the specifications for Class EA and Class FA permanent links and Category 6A and Category 7A components. The work on this document was started in early 2007 and is expected to be finalized during 2009.

As the requirements in Amendment 1 are more strict than the TIA/EIA document, itis also expected that the component requirements will be much stricter, therefore it will not be possible to use TIA/EIA compliant products to build a channel or link meeting the requirements of ISO/IEC 11801.

CENELEC EN 50173

CENELEC TC215 WG1, which is responsible for the EN 50173 series of documents, has decided to wait for the maturity of ISO/IEC 11801 before they start to create an amendment to EN 50173-1. The first amendment containing channels only for Class EA and Class FA is expected to be released in late 2008 and a 6 to 9 months delay is expected between Amendment 2 to ISO/IEC 11801 and the Amendment 2 to EN 50173-1.

The technical requirements for the amendments of EN 50173-1 will be specified as in ISO/IEC 11801.

8. Conclusion

With the release of IEEE 802.3an 10GBASE-T application standard in June 2006 and the following releases of ISO/IEC 11801 and TIA/EIA 568B.2-10 cabling standards, it must be considered to be relatively safe to invest in a future infrastructure enabling the telecommunication cabling system to carry 10GBASE-T signals.

There are still some small issues pending for component requirements in regard to the international standards, but these can be overcome by choosing a cabling system which is able to meet the channel specifications of Amendment 1 to ISO/IEC 11801.

The market forecast for server cards shows that within the next 5 years more than 99% of server cards will be 10GBASE-T, so it will be considered unwise to install less than a Class EA in your buildings, as this would limit the lifetime of the cabling infrastructure to 5 years. Cabling infrastructures should be considered to have a lifetime in excess of 10 years, which is also what all the cabling standards specifyin their introduction statements.

It still waits to be seen whether the dominant cabling will be a unshielded or an shielded infrastructure, as we now have a situation where the benefits of shielded cabling systems really benefit the application.

Last but not least it must be stated that IEEE 802.3 has commenced work on a new generation of Ethernet, the IEEE 802.3ba containing a 40G Ethernet and a 100G Ethernet version, as the first step for multimode and singlemode fibre optical cables, and a CX4 twinaxial solution, just as for 10GBASE in 2002, and studies have already started to investigate the feasibility of running 100G Ethernet over 100 meters of Category 7A / Class EA shielded cabling.

Notes:

AMP NETCONNECT Regional Headquarters:

North America
Greensboro, NC, USA
Ph: +1-800-553-0938
Fx: +1-717-986-7406

Latin America
Buenos Aires, Argentina
Ph: +54-11-4733-2200
Fx: +54-11-4733-2282

Europe
Kessel-Lo, Belgium
Ph: +32-16-35-2190
Fx: +32-16-35-2188

Mid East & Africa
Cergy-Pontoise, France
Ph: +33-1-3420-2122
Fx: +33-1-3420-2268

Asia
Hong Kong, China
Ph: +852-2735-1628
Fx: +852-2735-1625

Pacific
Sydney, Australia
Ph: +61-2-9407-2600
Fx: +61-2-9407-2519

AMP NETCONNECT in Europe, Middle East, Africa and India:

Austria – Vienna
Ph: +43-1-90560-1204
Fx: +43-1-90560-1270

Belgium – Kessel-Lo
Ph: +32-16-35-2190
Fx: +32-16-35-2188

Bulgaria – Sofia
Ph: +359-2-971-2152
Fx: +359-2-971-2153

Czech&Slovak Rep.–Kurim
Ph: +420-541-162-112
Fx: +420-541-162-223

Denmark – Glostrup
Ph: +45-70-15-52-00
Fx: +45-43-44-14-14

Egypt – Cairo
Ph: +20-2-2419-2334
Fx: +20-2-2417-7647

Finland – Helsinki
Ph: +358-95-12-34-20
Fx: +358-95-12-34-250

France – Cergy-Pontoise
Ph: +33-1-3420-2122
Fx: +33-1-3420-2268

Germany – Langen
Ph: +49-6103-709-1547
Fx: +49-6103-709-1219

Greece/Cyprus – Athens
Ph: +30-210-9370-396
Fx: +30-210-9370-655

Hungary – Budapest
Ph: +36-1-289-1007
Fx: +36-1-289-1010

India – Bangalore
Ph: +91-80-4011-5000
Fx: +91-80-4011-5030

Italy – Collegno (Torino)
Ph: +39-011-4012-111
Fx: +39-011-4012-268

Kazakhstan – Almaty
Ph: +7-327-244-5875
Fx: +7-327-244-5877

Lithuania – Vilnius
Ph: +370-5-213-1402
Fx: +370-5-213-1403

Netherlands – Den Bosch
Ph: +31-73-6246-246
Fx: +31-73-6246-958

Norway – Nesbru
Ph: +47-66-77-88-99
Fx: +47-66-77-88-55

Poland – Warsaw
Ph: +48-22-4576-700
Fx: +48-22-4576-720

Portugal – Evora
Ph: +351-961-377-331
Fx: +351-211-454-506

Romania – Bucharest
Ph: +40-21-311-3479
Fx: +40-21-312-0574

Russia – Moscow
Ph: +7-495-790-7902
Fx: +7-495-721-1894

Spain – Barcelona
Ph: +34-93-291-0330
Fx: +34-93-291-0608

Sweden – UpplandsVäsby
Ph: +46-8-5072-5000
Fx: +46-8-5072-5001

Switzerland – Steinach
Ph: +41-71-447-0-447
Fx: +41-71-447-0-423

Turkey – Istanbul
Ph: +90-212-281-8181
Fx: +90-212-281-8184

UK – Stanmore, Middx
Ph: +44-208-420-8140
Fx: +44-208-954-7467

Ukraine – Kiev
Ph: +380-44-206-2265
Fx: +380-44-206-2264

U.A.E. – Dubai
Ph: +971-4-321-0201
Fx: +971-4-321-6300

Lit#-MM/YY