10 Gigabit Ethernet

Siwaruk Siwamogsatham, siwamogs@ee.eng.ohio-state.edu

Abstract:

The next generation 10-Gigabit Ethernet Standard is currently under development and is expected to come out in the next few years. This paper highlights ongoing discussions regarding several aspects of 10-Gigabit Ethernet such as market requirements, implementation structures, layer interfaces, coding techniques, frame formats, media, devices, and etc.

See Also: Gigabit Ethernet | Gigabit Ethernet (Audio Recording of Lecture) | Gigabit Ethernet References | Books on Gigabit Ethernet | Optical Networking And Dense Wavelength Division Multiplexing (DWDM) | IP over DWDM
Other Reports on Recent Advances in Networking
Back to Raj Jain's Home Page

Table of Contents:

Introduction

Over the past several years, Ethernet has been the most popular choice of technology for local area networks (LAN). There are millions of Ethernet users worldwide and still counting. In 1998, the standard for 1-Gigabit Ethernet was released. It prompted a great deal of attention from users, especially many of those who were reluctant to adopt the expensive ATM technology for their LANs. Within a few years, 1-Gigabit Ethernet is likely to dominate the LAN markets.

As the demand for high-speed networks continues to grow, the need for a faster Ethernet technology is apparent. In March 1999, a working group was formed at IEEE 802.3 Higher Speed Study Group (HSSG) to develop a standard for 10-Gigabit Ethernet, which is expected to come out within the next few years. The preliminary objectives of the working group are listed below: [Grow99]

- Support 10 Gb/s Ethernet with about 2-3 times the cost of 1-Gigabit Ethernet
- Maintain the IEEE 802.3 Ethernet frame formats.
- Meet IEEE 802 Functional Requirements.
- Simple forwarding between all speeds.
- Maintain compatibility to IEEE 802.3x flow.
- Maintain minimum and maximum frame size of current IEEE 802.3 standard.
- Specify media independent interface.
- Full-duplex operation only.
- Speed-independent medium access control layer to support 10 Gb/s in LAN and about 10 Gb/s in MAN.
- Support star-wired LAN topologies.
- Support media selected from ISO/IEC 11801.
- Specify a family of physical layer which support a link distance of at least 200m on multi-mode fibers (MMF) and at least 3 km on single mode fibers (SMF).
- Support the existing cabling infrastructure as well as the new infrastructure.

This paper highlights ongoing discussions in the working group about several aspects of 10-Gigabit Ethernet. It should be pointed out here that this is still very early in the process of standard development. Many of the discussions are subject to change.

Back to Table of Contents

---

2. 10-Gigabit Ethernet Overview

10-Gigabit Ethernet is basically the faster-speed version of Ethernet. It will support the data rate of 10 Gb/s. It will offer similar benefits to those of the preceding Ethernet standard. However, it will not support the half-duplex operation mode. The potential applications and markets for 10-Gigabit Ethernet are enormous. There are broad groups of users who demand 10-Gigabit Ethernet, for example, enterprise users, universities, telecommunication carriers, and Internet service providers. Each market typically has different requirements for link span and cost.

2.1 Benefits of 10-Gigabit Ethernet

One of the main benefits of 10-Gigabit standard is that it offers a low-cost solution to solve the demands for bandwidth. Not only the cost of installation is low, but the cost of network maintenance and management is minimal as well. Management and maintenance for 10-Gigabit Ethernet may be done by local network administrators.

In addition to the cost reduction benefit, 10-Gigabit Ethernet may allow faster switching. Since 10-Gigabit Ethernet uses the same Ethernet format, it allows seamless integration of LAN, MAN, and WAN. There is no need for packet fragmentation, reassembling, or address translation, eliminating the need for routers that are much slower than switches.

10-Gigabit Ethernet also offers straightforward scalability (10/100/1000/10000 Mb/s). Upgrading to 10-Gigabit Ethernet should be simple since the upgrade paths are similar to those of 1-Gigabit Ethernet which should be well familiar by the time 10-Gigabit Ethernet standard is released. The main issues are that 10-Gigabit Ethernet is optimized for data and that it does not provide built-in quality of services. However, the quality of services may be provided in the higher layers.

2.2 Market Requirements

There are broad demands for 10-Gigabit Ethernet in the local area network (LAN), metropolitan area network (MAN), and the wide area network (WAN) markets. Each market typically has different requirements. Table 1. summarizes typical span requirements for different applications.

Applications	Typical Distance
Local Storage Area Networks	100 m
Remote Storage Area Networks	300 m/2 km/40 km
Disaster Recovery Facilities	300 m/2 km/40 km
Enterprise Networks	100 m
Enterprise Aggregation Facilities	550 m
Campus Backbone	10 km
Enterprise Backbone	300 m/ 2 km/ 40 km
Internet Service Providers	100 m
Internet Aggregation Facilities	300 m/2 km/40 km
Internet Backbone Facilities	300 m/2 km

Table 1. Distance Requirements [Bynum99]

In the LAN markets, applications typically include in-building computer servers, building-to-building clusters, and date centers. In this case, the distance requirement is relaxed, usually between 100 and 300 meters. But, the cost requirement is stringent. These users always look for cheep solutions to solve the demands for high-speed networking. The backward-compatibility is strongly desired because the users do not want the existing equipment to become obsolete. This makes 10-Gigabit Ethernet very attractive since most existing LANs are Ethernet-based.

In the medium-haul markets, applications usually include campus backbones, enterprise backbones, and storage area networks. In this case, the distance requirement is moderate, usually between 2 kms and 20 kms. The cabling infrastructure usually already exists. The technologies must operate over it. The initial cost is not so much on an issue. Normally, the users are willing to pay for the cost of installation but not the cost of network maintenance and management. The ease of technology is preferred. This give an edge to 10-Gigabit Ethernet over the only currently available 10-Gigabit SONET technology.

The WAN markets typically include Internet service providers and Internet backbone facilities. Most of the access points for long distance transport networks require the OC-192c data rate. So the key requirement for these markets is the compatibility with the existing OC-192c technologies. Mainly, the date rate should be 9.584640 Gb/s. Thus, the 10-Gigabit Ethernet stand shall specify a mechanism to accommodate the rate difference.

2.3 Protocol Layer

The Ethernet protocol basically implements the bottom two layers of the Open Systems Interconnection (OSI) 7-layer model, i.e., the data link and physical sublayers. Figure 1 depicts the typical Ethernet protocol stack and the relationship to the OSI model.

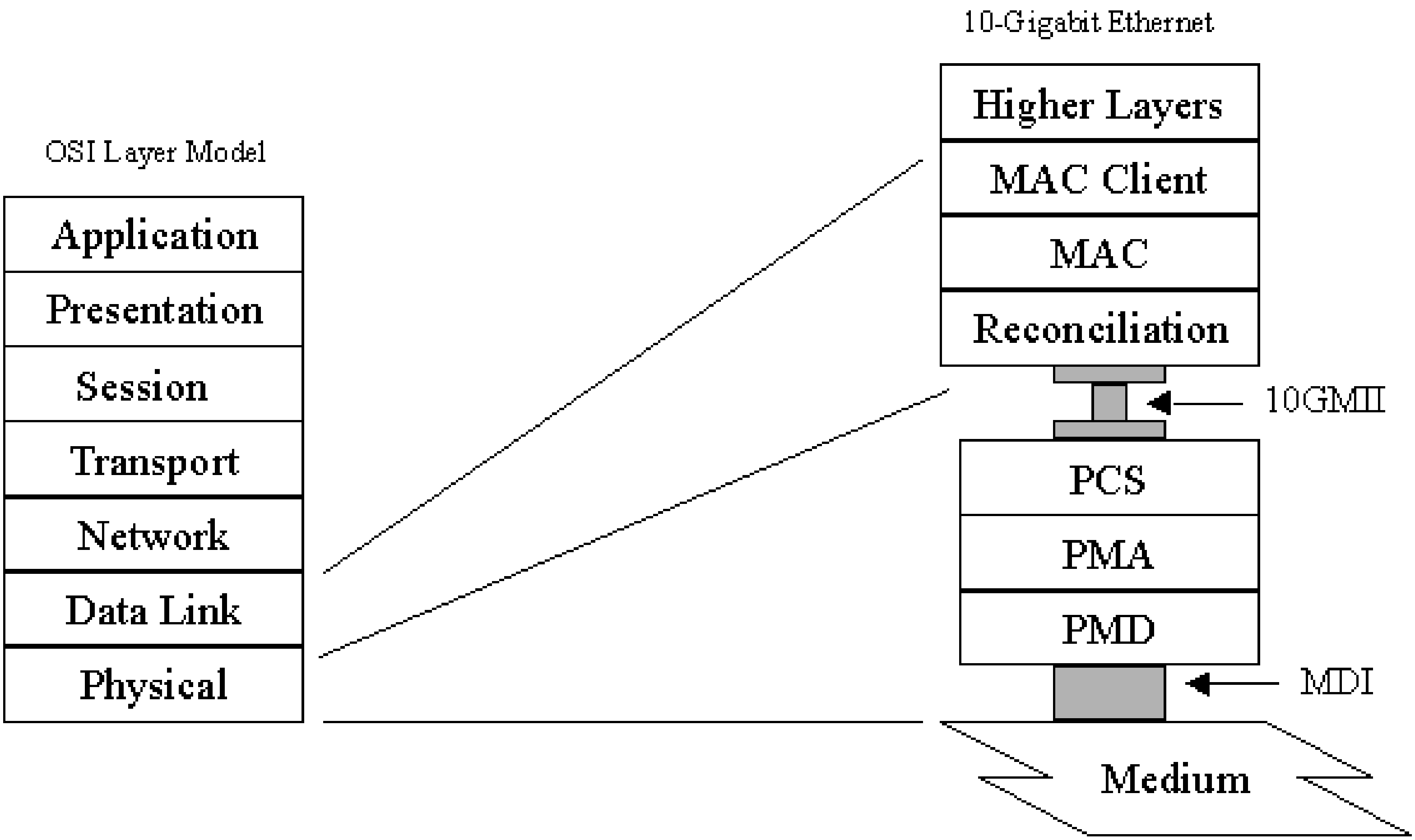

Figure 1. Ethernet Protocol Layer

Medium Access Control (MAC)

The media access control sublayer provides a logical connection between the MAC clients of itself and its peer station. It main responsibility is to initialize, control, and manage the connection with the peer station.

Reconciliation Sublayer

The reconciliation sublayer acts as a command translator. It maps the terminology and commands used in the MAC layer into electrical formats appropriate for the physical layer entities.

10GMII (10-Gigabit Media Independent Interface)

10GMII provides a standard interface between the MAC layer and the physical layer. It isolates the MAC layer and the physical layer, enabling the MAC layer to be used with various implementations of the physical layer.

PCS (Physical Coding Sublayer)

The PCS sublayer is responsible for coding and encoding data stream to and from the MAC layer. The default coding technique has not been defined. Several coding techniques will be discussed later in the paper.

PMA (Physical Medium Attachment)

The PMA sublayer is responsible for serialize code groups into bit stream suitable for serial bit-oriented physical devices and vice versa. Synchronization is also done for proper data decoding in this sublayer.

PMD (Physical Medium Dependent)

The PMD sublayer is responsible for signal transmission. The typical PMD functionality includes amplifier, modulation, and wave shaping. Different PMD devices may support different media.

MDI (Medium Dependent Interface)

MDI is referred a connector. It defines different connector types for different physical media and PMD devices.

Back to Table of Contents

---

3. 10-Gigabit Ethernet MAC Layer

The medium access control (MAC) layer of 10-Gigabit Ethernet will be similar to the MAC layer of previous Ethernet technologies. It will use the same Ethernet address and frame formats, but it will not support the half-duplex mode. It will support data rate of 10 Gb/s and lower, using pacing mechanism for rate adaptation and flow controls.

3.1 Full-Duplex Only

In the Ethernet standard, there are two modes of operation: half-duplex and full-duplex modes. The half-duplex mode has been defined since the original version of Ethernet. In this mode, data are transmitted using the popular Carrier-Sense Multiple Access/Collision Detection (CSMA/CD) protocol on a shared medium. Its simplicity contributed to the early success of the Ethernet standard. This mode of operation is so famous that many mistake the CSMA/CD protocol with the Ethernet standard. The main disadvantages of the half-duplex are the efficiency and distance limitation. In this mode, the link distance is limited by the minimum MAC frame size. This restriction reduces the efficiency drastically for high-rate transmission. For instance, the carrier extension technique is used to ensure the minimum frame size of 512 bytes in Gigabit Ethernet to achieve a reasonable link distance. At the transmission rate of 10 Gb/s, the half-duplex mode is not an attractive option. No market currently exists for the half-duplex operation at this rate of transmission. Most of the links at this rate are point-to-point over optical fibers. In this case, the full-duplex operation is the preferred option. It can be expected that the standard for 10-Gigabit Ethernet will specify only the full-duplex operation. In the full-duplex operation, there is no contention. The MAC layer entity can transmit whenever it wants, provided that its peer is ready to receive. The distance of the link is limited by the characteristic of the physical medium and devices, power budgets, and modulation. In this case, a desired topology can be achieved with the use of switches or distributed buffers.

3.2 MAC Frame Format

The key purpose the developing 10-Gigabit Ethernet standard is to use the same MAC frame format as specified in the preceding Ethernet standards. This will allow a seamless integration of the 10-Gigabit Ethernet with the existing Ethernet networks. There is no need for fragmentation/reassembling and address translation, implying faster switching. Since the full-duplex operation is used, the link distance does not affect the MAC frame size. The minimum MAC frame size will be made equal to 64 octets as specified in the previous Ethernet standards. Carrier extension is not needed. The MAC frame format is depicted in Figure 2.

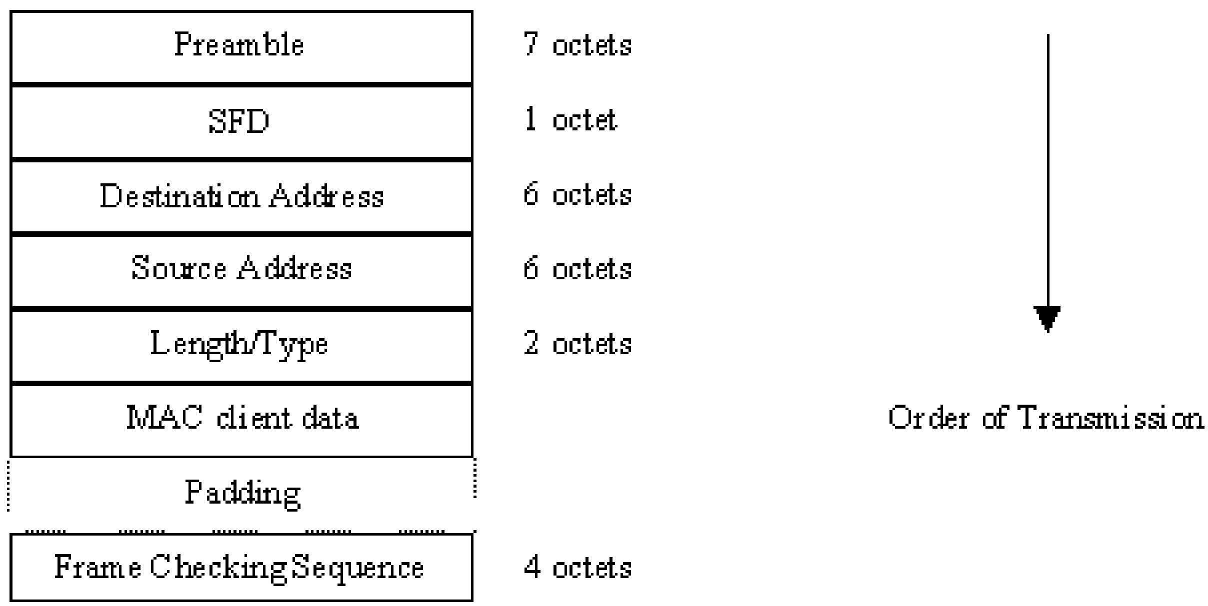

Figure 2. Ethernet Frame Format

The Ethernet frame format consists of the following fields:

- **Preamble**. A 7-octet a preamble pattern of alternating 0's and 1's that is used to allow receiver timing synchronization to reach a steady state.
- **Start frame delimiter (SFD)**.The SFD field is the sequence 10101011, used to indicate a start of frame.
- **Address fields**. Each MAC frame contains .the destination and source addresses. Each address is 48 bits long. The first of which is used to identify the address as an individual address (0) or a group address (1). The second of which is used to indicate whether the address is locally (1) or globally (0) defined.
- **Length/Type** If the number is less than the maximum valid frame size, it indicates the length of the MAC client data. If the number is greater than or equal to 1536 decimal, it represents the type of the MAC client protocol.
- **Data and padding**. Padding is optional. It is only necessary when the data packet is smaller than 38 octets to ensure the minimum frame size of 64 octets as specified in the existing standards.
- **Frame checking sequence (FCS)**. The FCS field contains a 32-bit cyclic redundancy check

(CRC) value computed from all fields except the preamble, SFD, and CRC. The encoding is defined by the following generating polynomial:

$$G(x) = x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^8+x^7+x^5+x^4+x^2 x + 1$$

3.3 Data Rate

Specifying data rata for 10-Gigabit Ethernet is no simple task. After several months of discussions, the data issue remains undecided. People in the LAN business want the data rate to be 10 Gb/s so that a 10-Gigabit Ethernet switch can support exactly ten 1-Gigabit Ethernet ports. People in the WAN business, on the other hand, want the 9.584640 Gb/s data rate so that it is compatible with the OC-192 Rate.

The solution the working group is leaning toward is to support both rates. This can be done by specifying the data rate at 10 Gb/s and utilizing pacing mechanism to accomodate the slower data rates. What might be an issue in this solution is that it requires a device with a large buffer to bridge the two rates. It should be noted that if the data rate is specified at 9.584640 Gb/s, it is not possible to support the 10 Gb/s data rate. Therefore, it is more likely that the 10 Gb/s data rate will win the battle.

3.4 Pacing Mechanism

The pacing mechanism allows the MAC layer entity to support transmission rates, for instance, 1 Gb/s or 10 Gb/s for LAN and 9.584640 Gb/s for WAN. To achieve this, the MAC layer entity shall have an ability to pause data transmission for an appropriate period of time to provide a flow control or rate adaptation. Two techniques for pacing mechanism are under consideration. The first is the word-by-word hold mechanism and the second is the Inter-Frame GAP (IFG) stretch technique. In the word-by-word technique, the MAC layer entity pauses sending a 32-bit word of data for a pre-specified period of time upon a request from the physical layer. In the IFG stretch technique, the IFG is extended for a pre-defined period of time with or without a request from the physical layer. The main disadvantage of the IFG stretch technique is that a large data buffer is required because the algorithm operates between frames. The word-by-word hold mechanism seems to obtain a stronger support from the working group. The main advantages of the word-by-word mechanism are that it can support any encoding techniques, it does not need a large data buffer to hold multiple MAC frames, and the buffer size is independent of link speed.

Back to Table of Contents

---

4. 10-Gigabit Ethernet Physical Layer

Most of the ongoing discussions in the working group are related to the physical layer. The main issues include 10-Gigabit Media Independent Interface, parallel vs. serial architectures, wavelength division multiplexing (WDM) vs. parallel optic, coding techniques, devices, media, and so on.

4.1 10-Gigabit Media Independent Interface (10GMII)

10-Gigabit Media Independent Interface (10GMII) provides the interface between the MAC layer and the physical layer. It allows the MAC layer to support various physical layer variations. The current proposal for 10GMII [Frazier99] is depicted in Figure 3.

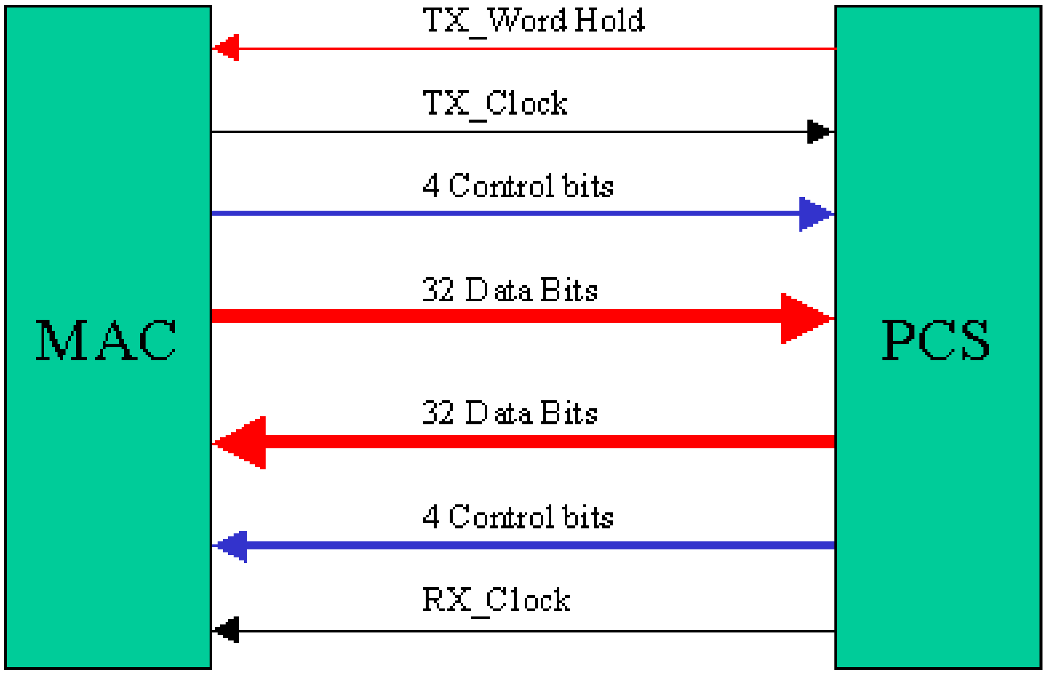

Figure 3. 10GMII Proposal

The TX_word_hold line is provided to support word-oriented pacing mechanism. The 32-bit data paths are provided for transmit and receive functions each with 4 control bits (one per byte). The control bit is set to "1" for delimiters and special characters and "0" for data. Delimiters and special characters are determined from the 8-bit data value when the control bit is set to "0". The delimiter and special characters include:

- **IDLE** which is signaled during the inter-packet gap and when there is no data to send
- **SOP** which is signaled at the start of each packet
- **EOP** which is signaled at the end of each packet
- **ERROR** which is signaled when an error is detected in the received signal or when an error needs to be put the to the transmitted signal.

These delimiter and special characters enables a proper synchronization for multiplexing and demultiplexing operations. It should be noted that the interface could also be scaled in speed and width. For example, an 8-bit data path with 1 control bit may be used at 4 times faster clock rate. In this way, the delimiter and special characters remain unchanged. Thus, it can support both serial and parallel implementations of the PCS sublayer.

4.2 Physical Layer Architecture

There are two structures for the physical layer implementation of 10-Gigabit Ethernet: the serial solution and parallel solution. The serial solution uses one high-speed (10 Gb/s) PCS/PMA/PMD circuit block and the parallel solution uses multiple PCS/PMA/PMD circuit blocks at lower speed. The two solutions have different advantages and disadvantages which are discussed next.

4.2.1 Serial Implementation

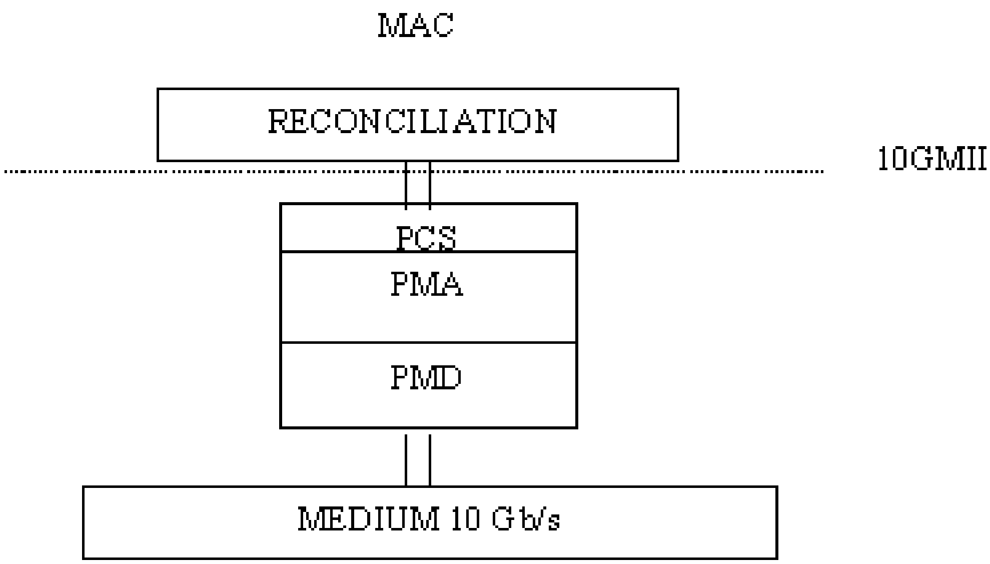

Figure 4. Serial Physical Layer Implementation

In the serial implementation, there is one physical channel operating at 10 Gb/s as depicted above. The operation is straightforward. For transmission, the reconciliation module passes the signals, corresponding to the MAC data, word by word to the PCS module. The PCS module then encodes the signals with a pre-defined coding technique and passes the encoded signal to the PMA module. The PMA module then serializes the encoded signals and passes the stream to the PMD module. The PMD module transmits the signal stream over the fiber at 10 Gb/s. For receiving, the process is reverse.

The main advantage of the serial architecture is that the transmitting/receiving operation is straightforward. It does not require a complicated multiplexing/demultiplexing that is needed in the parallel implementation. Thus, the timing jitter requirement is more relaxed. It also requires only one fiber channel and one set of laser equipment. However, the main disadvantage is the need of expensive high-speed logic circuits technology such as SiGe and laser equipment. To reduce the transmission rate, higher-rate coding techniques such as PAM-5 and MB810 may be used. In that case, only one low-cost laser unit may be needed. More research should be done to investigate this benefit.

There are technologies, for example the 10G-SONET/OC-192, which currently support 10 Gb/s operation. The technologies from these existing standards may be borrowed to aid the 10G Ethernet serial implementation. Therefore, for those who want a quick solution and are less concerned about costs, the serial architecture should be a good fit. The strong supporter for the serial implementation is Lucent who has demonstrated a serial transmission at 10 Gb/s and 12.5 Gb/s over the new high-bandwidth MMF up to 300m.

4.2.2 Parallel Implementation

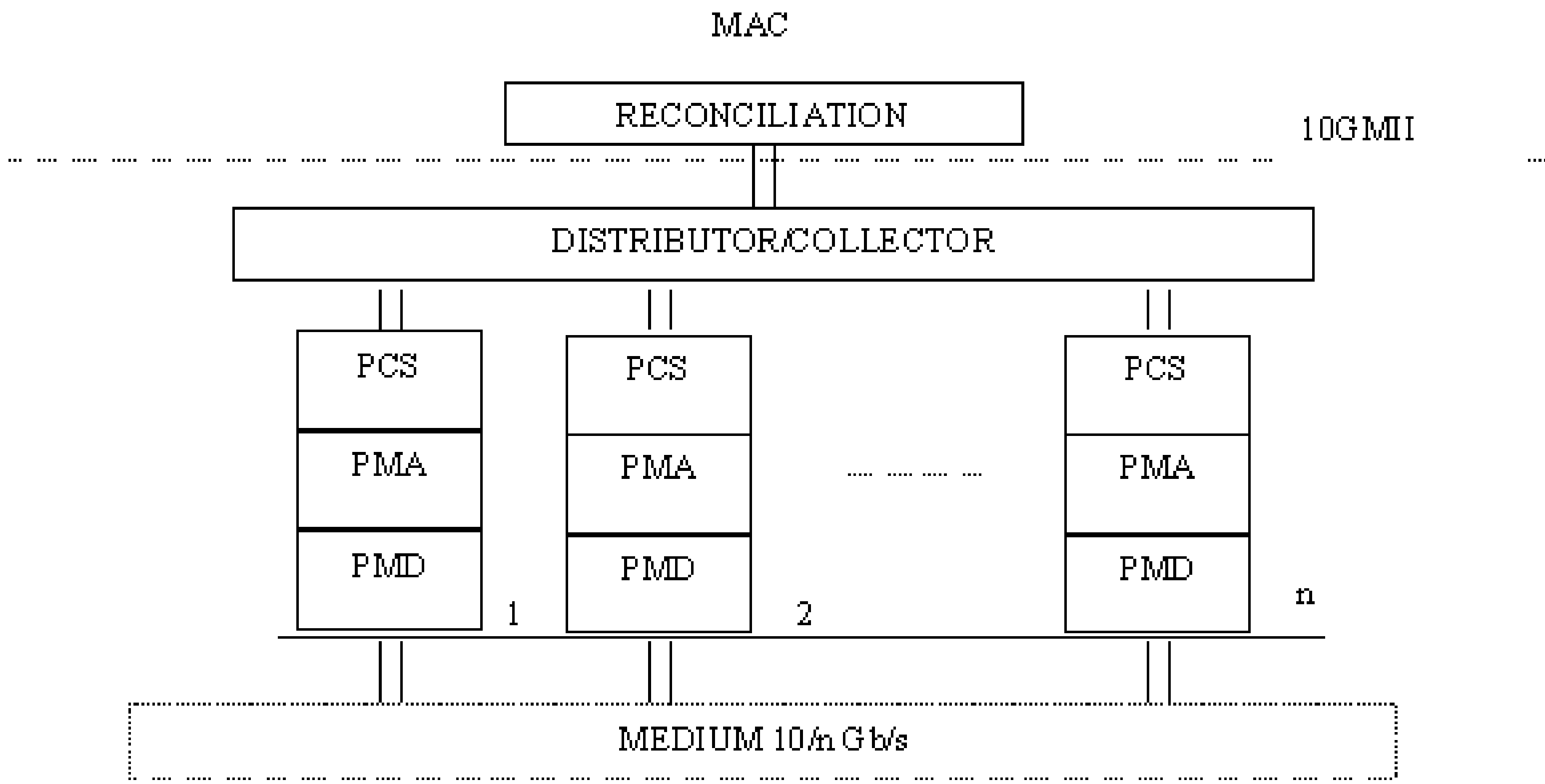

Figure 5. Parallel Physical Layer Implementation

In parallel implementation, there are multiple physical channels, say n sub-channels that may be implemented by using parallel cables or WDM multiplexing. For transmission, the distributor multiplexes the data (frames and idles) accepted from the MAC layer into n streams ("mini-frames", "mini-idle") in the Round-Robin motion. Each stream is given to each PCS module. Each PCS module encodes the received stream and passes it to each PMA module for serialization. After serialization, each PMD module transmits each serialized data stream at the rate of 10/n Gb/s. For receiving, the reserve process is done.

The main advantage of the parallel implementation is that the operating rate in the PCS/PMA modules is reduced, which enables cheaper devices (CMOS/Bipolar) to be used. The disadvantages are the need of distributor/collector module that may be sensitive to timing jitters, and the usage of multiple sets of logic circuits and laser equipment. There are two techniques to achieve multiple channels, one of which is the parallel cabling and the other is the WDM.

For the parallel solution to achieve a total rate of 10 Gb/s, 10x 1 or 10x 1.25 Gb/s fibers may be used. Since the 1-Gigabit is available now, this upgrade is already possible. However, the cost of this solution is 10 times of the 1-G Ethernet. This violates the economic objective of 10-G Ethernet, making it not an attractive solution now. But when 2.5 Gb/s equipment becomes available, only 4 parallel fibers are needed. Then, this technique will be very attractive for short-haul (<200 m) applications where the cost of expensive WDM optical equipment may outscore the cost of fiber cabling. However, the parallel cabling solution does not apply to the existing infrastructure. In contrast, the WDM solution can be used over the existing cables but the big issue is the expensive WDM optical equipment.

4.3 Coding Techniques

This section discusses several coding techniques that might be adopted in the 10-G Ethernet standard.

4.3.1 8B/10B

This code was invented by IBM for low-cost devices and implementations. It has an excellent DC balance property with a maximum run-length of 5 and a good transition density, which simplifies the requirements on bandwidths and clock jitters for device circuits. With an input/output bit rate of 8/10, this code can improve transmission reliability by some degree. It can also provide built-in special characters for commands, synchronization, and delineation. The encoding/decoding algorithm for this code is simple and can be implemented in low-cost hardware. Apart from the aforementioned benefits, this code has another key advantage. It was adopted in the recently released, 1-Gigabit Ethernet standard (IEEE 802.3z), making it very attractive to the 10-Gigabit Ethernet. The use of 8B/10B coding will make 1-Gigabit Ethernet and 10-Gigabit Ethernet truly compatible, enabling a seamless integrated Ethernet network with less cost, effort, and time for technology migration. This advantage will likely give an edge to the 8B/10B coding in the early deployment of the 10-G Ethernet standard for LAN/MAN. The main drawback of this coding technique is the 25 % overhead and less bandwidth efficiency. That is, a line rate of 12.5 Gb/s is needed to implement a transmission rate of 10 Gb/s with the 8B/10B encoding. For a serial implementation, this is a significant disadvantage because currently there are a larger number of devices that can support 10 Gb/s, but very few at 12.5 Gb/s.

4.3.2 Scrambling

The scrambled encoding has its place in SONET/WAN applications. It has virtually no overhead and better bandwidth efficiency, allowing a lower line rate for extended reach. The implementation of this code is also simple and can be done in hardware. However, its maximum run-length is non-deterministic, it has no guarantee of DC balance, and no built-in special characters. The cost of devices with this encoding scheme is typically higher, which could make this technique less attractive for a LAN unless an extended reach is needed. One may argue that using scrambled encoding scheme will enable a seamless Ethernet in LAN, MAN, and WAN, plus signaling technologies can be directly borrowed from SONET. However, it may lack a physical compatibility with IEEE 802.3z standard. In practice, a compromise has to be made. A bridge may be needed somewhere between LAN and WAN in order to handle the difference of line encoding techniques and other attributes. The standard is likely to define different specifications for LAN and MAN/WAN. The 8B/10B encoding will likely be used LAN/MAN while the scrambled encoding will likely be specified in the WAN.

4.3.3 MB810

This coding technique is recently proposed by ChanGoo Lee, et al [Lee99]. It has similar advantages and disadvantages as those in 8B/10B. The key benefit of this code is the 50 % improvement of bandwidth efficiency over the 8B/10B code. If this claim can be strongly proven, the line rate may be reduced in half. That is, only 5 Gbaud is needed for a serial implementation of 10-Gigabit Ethernet. This benefit sounds very promising. However, until the technologies supporting the MB810 encoding are proven, a main role of this coding technique is not expected especially in the early stage of the standard development.

4.3.4 PAM-5

This coding technique was adopted in the1000-Base-T standard. It employs multi-level amplitude signaling to increase the number of bits per baud. In the twisted-pair lines, the PAM-5 encoding can achieve 2 bits per baud with a 3 dB coding gain, yielding a significantly lower line rate. For fiber channels, however, more works need to be done to study the impact of the ISI, SNR, and non-linearity penalty on a multi-level amplitude signaling. The PAM-5 signaling may not travel in the fiber very far, and the current devices may not support this signaling. Until these concerns are rigorously answered, the PAM-5 encoding will not compete with the 8B/10B and scrambled encoding.

4.3.5 16B/18B

This code has similar benefits as those of the 8B/10B but with fewer overheads, 12.5% compared with 25% overhead for the 8B/10B. However, it is not compatible with the physical layer of the IEEE 802.3z standard nor SONET. This code is less attractive.

4.3.6 Forward Error Correction

The Forward Error Correction (FEC) code has a capability of correcting some corrupted received data. It effectively lowers the bit error rate (BER) by a great amount with an expense of some additional overhead, for example, with 6% overhead an FEC code can achieve 10-14 BER with 10-4 BER input. The poplar FEC codes are the BCH codes and the RS codes. It can be used in conjunction with the 8B/10B or scrambled encoding. This technique is useful in applications, such as long-haul networks, where coding advantage is strongly needed. In 10G-SONET, an FEC encoding based on a BCH code is likely to be adopted. The same trend may carry over the 10-G Ethernet for WANs.

4.4 Lasers

An essential PMD component for high-speed transmission is laser. There are sereral types of lasers. The common ones are the Fabry-Perot (F-P) Laser, Vertical-Cavity Surface-Emitting Laser (VCSEL), and Distributed-Feedback (DFB) Laser.

4.4.1 Fabry-Perot Laser

The Fabry-Perot (F-P) laser is simple lost cost multi-mode laser. It is optimized for single mode fibers but it can also operate over multi-mode fibers. The typical operating wavelength is in the 1300-nm range. For this type of optical source, the distance limitation is due to dispersion and mode-partition noise. The Fabry-Perot laser technology is matured and proven. The typical applications for Fabry-Perot lasers are for campus LAN and MAN networks.

4.4.2 Vertical-Cavity Surface-Emitting Laser

The VCSEL laser is traditionally a low-cost solution for 850-nm application. It can operate on both multi-mode and single mode fibers. For this type of source, the link distance is quite limited. The recent technology (LW-VCSEL) allows VSCEL lasers to operate in the 1300-nm range improving the distance limitation drastically.

4.4.3 Distributed-Feedback Laser

The Distributed-Feedback laser utilizes distributed resonators to suppress multi-mode source. It has high bandwidth-distance product and typically operates over the 1300-nm wavelength band on single-mode and multi-mode fibers, and 1550-nm band on single mode fibers. It allows both direct and external modulation. The distance limitation is typically due to attenuation loss for the 1300-nm band and dispersion for the 1550-nm band. Thermal cooling and isolators may be used enhance performance.

4.5 Physical Media

The physical media for high-speed transmission are typically optical fibers. Table 2 describes various types of fibers and the corresponding estimated distance limitation for different types of optical sources and line rates.

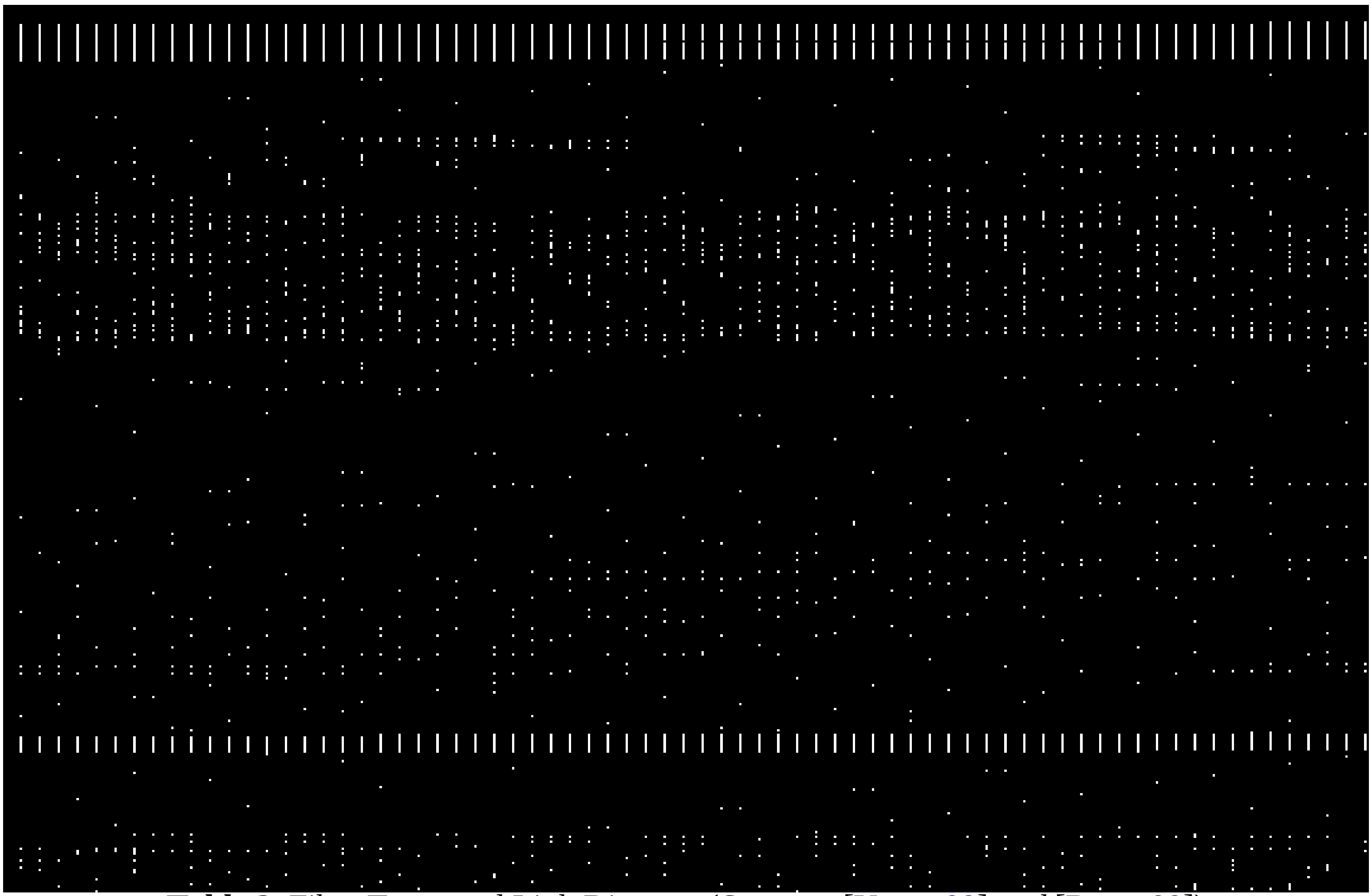

Table 2. Fiber Types and Link Distance (*Sources:* [Young99] *and* [Doorn99])

4.5.1 62.5-um Multi-mode Fiber

The 62.5-um multi-mode fiber is the cheapest option among the applicable choices of fibers. Most of the existing fiber infrastructures for links up to 300 meters are 62.5-um multi-mode fibers. It typically supports operations in the 800 nm and 1300 nm wavelength bands. The performance of this type of fiber is typically limited at about 200 MHz*km limiting the link distance to less than 50 meters for a line rate about 10 GBaud. In practice, a parallel solution is needed to achieve the total rate of about 10 GBaud over the existing 62.5-um multi-mode infrastructures. An example is the WDM system demonstrated by HP, using four wavelengths (1280,1300,1320,1340 nm) with uncooled DFB sources over the 62.5 um multi-mode fiber up to 300 meters. The typical applications of the 62.5 multi-mode fibers are in the in-building and building-to-building networks.

4.5.2 50-um Multi-mode Fiber

The traditional 50-um multi-mode fiber has slight better performance than the 62.5-um multi-mode fiber. In this case, link distance is limited to less than 65 meters at about 10 Gbaud line rate. However, the new enhanced 50-um multi-mode fibers such as ZETA by Lucent have much better performance. Standardization of this new type of fibers is ongoing. For this type of fibers, the bandwidth-distance product may be as high as 2000 MHz*km, enabling a line rate of 12.5 GBaud for link distance up to 300 meters. This implies the feasibility of a low-cost serial transmission solution. The technology was demonstrated by Lucent.

4.5.3 Single Mode Fiber

The single-mode fiber has smaller core than the multi-mode fiber, enabling signals to travel much longer distance. At the rate line about 10 GBaud, the link can be as long as 40 km. In practice, the single-mode fibers are suitable for LAN backbones, MAN, and WAN.

4.5.4 Copper Wires

In addition to optical fibers, copper wires may be used. In 1-Gigabit Ethernet, there is an option for copper-based physical channels. Copper wires are much cheaper than fibers, but the link distance is much shorter. A typically application may be the very short-haul switch/server jumper for distance up to 20 meters. Since the cost of copper wire installation is not an issue, the parallel-cable solution might be attractive for this case. In the early version of 10-Gigabit Ethernet, however, the copper wire option is not likely to be considered since it might slow down the standard development.

4.6 Distance

In 10-Gigabit Ethernet, the link span is limited solely by the capability of devices and media since only the full-duplex mode is supported. The issue is basically to define the right technology for the right market. As mentioned earlier, there are different types of markets: very short-haul, short-haul, medium-haul, and long-haul. Each market has different distance and cost requirements. Therefore, the solution that optimizes for all types of markets is not optimum. Instead, there should be different solutions optimized for different types of markets. In other words, different physical layer specifications shall be defined for different distance limits. It is noted that the similar thing was done in the past. The tentative physical layer specifications are listed in Table 3.

Interface	Type of media	Estimated Distance
10GBase-SX	Multi-Mode Fiber	100-300 m
10GBase-LX	Single Mode and Multi-Mode Fiber	5 - 15 km

10GBase-EX	Single Mode Fiber	> 40 km
10GBase-CX	Copper Wire	< 20 m

Table 3. Tentative Interfaces for 10-Gigabit Ethernet (*Source:* http://www.10gigabit-ethernet.com)

4.7 Bit Error Rate

To achieve a reliable transmission at 10 Gb/s data rate, the bit error rate (BER) needs to be reasonable small. A simple calculation may be done to roughly show how often errors may occur for a certain BER.

BER	# Errors for 10 Gb/s rate
10^(-10)	An error in every second
10^(-11)	An error in every 10 seconds
10^(-12)	An error in every 1 minute and 40 seconds
10^(-13)	An error in every 16 minutes 40 seconds

Table 4. Number of Errors for Given BER

As seen in Table 4, BER should be less than 10^(-12) to have a reasonably reliable transmission.

Back to Table of Contents

Summary

The huge of demand for high-speed networking has stirred development of the next generation 10-Gigabit Ethernet. The 10-Gigabit Ethernet standard will support the data rate of 10 Gb/s, which is 10 times faster than the current transmission rate of 1-Gigabit Ethernet, but the cost is targeted around 2 to 3 times the cost of the current 1-Gigabit Ethernet technology. It will not support the half-duplex operation, but it will maintain the compatibility with the preceding Ethernet technologies by using the same Ethernet frame format, enabling seamless integration among the existing networks and the new technology. 10-Gigabit Ethernet may be used in LAN, MAN, and WAN, implying technology convergence and faster switching.

In addition to the data rate of 10 Gb/s,10-Gigabit Ethernet shall be able to accommodate slower date rates such as 9.584640 Gb/s (OC-192). This is likely to be done by using the word-by-word pacing mechanism via the 10GMII interface. The current proposal for 10GMII uses 32-bit data paths with 4 control bits (one control bit per one data byte) and a TX_hold line. This structure provides scalability and can support various variations of the physical layer implementation. The discussions regarding issues in the physical layer are still ongoing. The standard may provide different specifications for different applications. The main considerations include the implementation architectures (serial, parallel, or WDM), coding techniques, supporting technologies, and media.

The 10 Gb/s serial transmission solution appears to be the easiest and lowest-cost option. The recent technology demonstration by Lucent validates the 10 Gb/s serial transmission using 850-nm VCSEL sources on the enhanced multi-mode fibers over the link longer than 300 meters. However, it cannot support the existing multi-mode fiber infrastructure. On the other hand, the WWDM solution supports the existing multi-mode fiber infrastructure but the implementation is more complicated and the devices are still more expensive. The 4 x 3.125 GBaud WDM solution using 1300-nm uncooled DFB sources on the existing 62.5-um multi-mode fiber can reach the link distance up to 300 meters. In addition to the WDM solution, the parallel fiber option may be suitable for short-haul applications such as computer rooms. In addition to fibers, copper wires may be used for a very short link (10-20 meters) but may not be considered in the early version of the standard.

The coding issue still remains opened. Several coding techniques are considered. The trend is that the 8B/10B encoding technique may be specified for the LAN rate Ethernet because it is compatible with 1-Gigabit Ethernet, while the SONET-like scrambled encoding techniques may be specified for the WAN rate Ethernet because it is compatible with SONET. These coding techniques are proven and well-understood, but the problem with the 8B/10B is the large 20% overhead. Other coding techniques such as 15B/16B, MB810, and PAM-5 are considered as well.

In short, 10-Gigabit Ethernet will be a low-cost solution for high-speed and reliable data networking and it will the dominate the LAN, MAN, and WAN markets in the near future. Let's go 10-Gigabit Ethernet!

Back to Table of Contents

References

[Grow99] Bob Grow, "Objectives for High-Speed 802.3," http://grouper.ieee.org/groups/802/3/10G_study/public/june99/grow_1_0699.pdf

[Young99] L. Young, V. Bhatt, and E. Cornerjo et. al., "Preliminary Serial PMD Proposal for HSSG," http://grouper.ieee.org/groups/802/3/10G_study/public/nov99/cornejo_1_1199.pdf

[Doorn99] S. V. Doorn, "10 Gig Status and Technology," http://grouper.ieee.org/groups/802/3/10G_study/public/march99/van_doorn_1_03991.pdf

[Muller99] Shimon Muller, "Architecture for 10 GbE Standard," http://grouper.ieee.org/groups/802/3/10G_study/public/june99/muller_1_0699.pdf

[Salzman99] Mike Salzman, "10 GbE requirements & architecture," http://grouper.ieee.org/groups/802/3/10G_study/public/june99/salzman_1_0699.pdf

[Dolfi99] Dave Dolfi, "10 Gb/s Wide WDM Transceivers (SpectraLAN) at 850 nm and 1300 nm," http://grouper.ieee.org/groups/802/3/10G_study/public/march99/dolfi_1_0399.pdf

[Frazier99] Howard Frazier, "10Gig MII Update," http://grouper.ieee.org/groups/802/3/10G_study/public/nov99/frazier_1_1199.pdf

[Ewen99] John Ewen, "10 Gb/s Physical Layer Options and Coding issues," http://grouper.ieee.org/groups/802/3/10G_study/public/june99/ewen_1_0699.pdf

[Musk99] Bob Musk and Alan Flatman, "In-Premises Optical Fibre Deployment," http://grouper.ieee.org/groups/802/3/10G_study/public/july99/flatman_1_0799.pdf

[Frazier99b] Howard Frazier, "Two Physical Layers for 10 Gigabit Ethernet," http://grouper.ieee.org/groups/802/3/10G_study/public/sept99/frazier_1_0999.pdf

[Bynum99] Roy Bynum, "10GbE Implementation Architectures," http://grouper.ieee.org/groups/802/3/10G_study/public/sept99/bynum_1_0999.pdf

[Giaretta99] Giorgio Giaretta, R.Michalzik, and P.Kolesar. "New MMF, how far can we go?," http://grouper.ieee.org/groups/802/3/10G_study/public/nov99/giaretta_1_1199.pdf

[Dugan99] Richard Dugan, "PCS/PMA interfaces for 10 Gb/s Ethernet,"http://grouper.ieee.org/groups/802/3/10G_study/public/june99/dugan_1_0699.pdf

[Lee99] ChanGoo Lee, D. Y. Kim, H. W. Jung, and H. H. Lee "MB810 Line Code for 10GbE," http://grouper.ieee.org/groups/802/3/10G_study/public/july99/kim_1_0799.pdf

[Dudley99] Thomas Dudley "Components and costs considerations for 10 Gbits/sec in the enterprise backbone," http://pennwell.shore.net/lightwave/archives/feat/feat0399.html

Links

- 10 Gigabit Ethernet Resource Site
- IEEE802.3 Higher Speed Study Group

Back to Table of Contents

List of Acronyms

10GMII- 10-Gigabit Media Independent Interface
ATM- Asynchronous Transfer Mode
BER- Bit Error Rate
CRC- Cyclic Redundancy Check
CSMA/CD- Carrier Sense Multiple Access/Collision Detection
DFB- Distributed-Feedback
FCS- Frame Checking Sequence
F-P- Fabry-Perot
Gb/s- Giga bits per second

Gbaud- Giga Baud
IFG- Inter-Frame Gap
km- Kilometer
LAN- Local Area Networks
m - Meter
MAC- Medium Access Control
MAN- Metropolitan Area Networks
Mb/s - Mega bits per second
MDI- Medium Dependent Interface
MMF- Mulit-Mode Fibers
nm - Nanometer
OC - Optical Carrier
OSI- Open Systems Interconnection
PCS- Physical Coding Sublayer
PMA- Physical Medium Attachment
PMD- Physical Medium Dependent
SFD- Start of Frame
SMF- Single Mode Fibers
VCSEL- Vertical-Cavity Surface-Emitting Laser
WAN- Wide Area Networks
WDM- Wavelength Division Multiplexing
WWDM- Wide Wavelength Division Multiplexing

Back to Table of Contents

Last Modified: November 23,1999.
Note: This paper is available on-line at http://www.cis.ohio-state.edu/~jain/cis788-99/10gbe/index.html