

Project: IEEE 802.3z Gigabit Ethernet Task Force

Source: Steve Dabecki (steve_dabecki@pmc-sierra.bc.ca)
Barry Hagglund (barry_hagglund@pmc-sierra.bc.ca)
Richard Cam (richard_cam@pmc-sierra.bc.ca)
Vern Little (vern_little@pmc-sierra.bc.ca)
Iain Verigin (iain_verigin@pmc-sierra.bc.ca)

PMC-Sierra Inc
105-8555 Baxter Place
Burnaby
BC
Canada
V5A 4V7

Tel: 604.415.6000

Web: <http://www.pmc-sierra.com>

Title: Physical Layer Proposal for 1 Gbit/s Full / Half Duplex Ethernet
Transmission over Single (4-pair) or Dual (8-pair) UTP-5 Cable.

Issue: 2.0

Date: November 11-14, 1996

Abstract

The extension of the current IEEE802.3 specifications to support a bit rate of 1Gbps is currently under investigation by the HSSG. These bit rates are regarded to be essential to support the high bandwidth requirements of mixed 10BASE and 100BASE networks, and many applications requiring a dedicated high speed link to a server. The HSSG has already made excellent progress in the way of defining 1000BASE for fibre. This contribution outlines an approach for supporting 1Gbps, either half or full-duplex, over single (4-pair) or dual (8-pair) unshielded twisted pair category 5 (UTP-5) cable.

Notice

This contribution has been prepared to assist the IEEE802.3z HSSG. This document is offered to the IEEE802.3z HSSG as a basis for discussion and is not a binding proposal on PMC-Sierra, Inc. or any other company. The statements are subject to change in form and/or content after further study. Specifically, PMC-Sierra, Inc. reserve the right to add to, amend or modify the statements contained herein.

Physical Layer Proposal
for 1 Gbit/s Full / Half Duplex Ethernet
Transmission over Single (4-pair) or
Dual (8-pair) UTP-5 Cable

Version 2.0

November 1996

1. Introduction.....	4
2. 1000BASE-TX Functional Overview.....	6
3. Existing Ethernet Signal Coding	9
3.1 4B5B and 8B10B Coding	9
3.2 Serial/Parallel Conversion (SERDES)	10
4. Signal Coding with 4LZS for 1000Base-TX.....	11
4.1 Escape Sequence Overview	11
4.2 Pattern Substitution Escape Sequences.....	12
4.3 Escape Sequence and PCS Layer Streams.....	12
4.4 Escape Sequence Mappings	14
4.5 Physical Layer Streams for Single/Dual Cables.....	17
4.6 Frame Scrambling	18
4.7 Comparison of 4LZS against NRZ.....	20
4.7.1 Theoretical Analysis of 2B1Q against NRZ.....	21
4.8 Receiver Line Equalization	22
4.9 Single or Dual Cable Operation.....	23
5. Gigabit Media Independent Interface (GMII)	24
5.1 GMII Functional Overview.....	24
5.2 GMII Signals and Timing Requirements.....	26
5.3 GMII Functional Timing	27
5.4 GMII AC Timing.....	28
5.5 GMII Management Interface	28
6. Simulation Model and Results.....	29
6.1 Overview	29
7. Hardware Evaluation and Results.....	31
7.1 Overview	31
7.2 Test Configuration.....	32
7.3 Test Procedure	35
7.4 Test Results	35
8. EMC Testing.....	42
8.1 Overview	42
8.2 Test Setup	43
8.3 Test Procedure	44
8.4 Test Results.....	44
9. Conclusion.....	49
10. References	50

1. Introduction

This contribution describes the Physical Layer components required to support transmission of Ethernet frames at 1Gbps full duplex data rate carried over single or dual UTP-5 cable. Transmission over a standard 4 pair UTP-5 bundled cable would provide a reach of 50m, and transmission over 2 such standard bundled cables would provide a reach of 100m. The method is suitable for both full and half duplex modes of operation.

Operation with single and dual cables is shown in Fig 1. below:

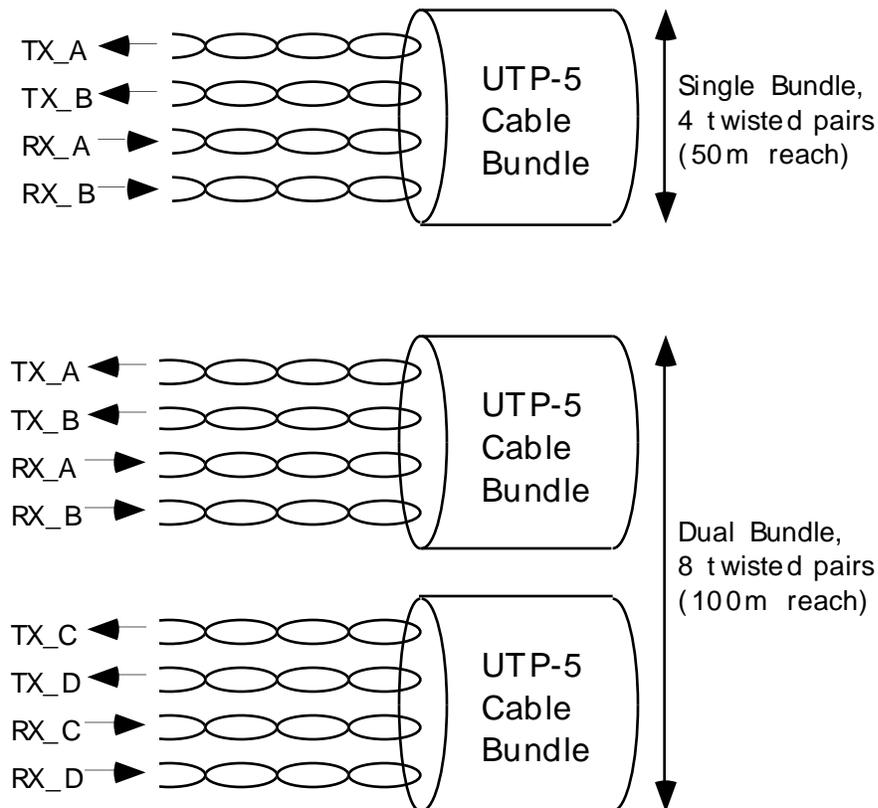


Fig 1. Operation with Standard Single or Dual UTP-5 Cable Bundles

In full duplex mode, a single cable would use 2 pairs for transmit and 2 pairs for receive, each pair carrying 500Mbps of data. Alternatively, dual cables could be used to double the reach to 100m, 4 pairs being used for transmit and 4 pairs for receive, each pair carrying 250Mbps. With full duplex there is no need for higher layers to support the CSMA/CD protocol as the connection would be point to point.

In half-duplex mode, a similar transmit and receive scenario would exist with either single or dual cables, but the higher layer would need to implement the CSMA/CD protocol.

There are numerous problems in trying to transmit data at these rates using only unshielded twisted pairs (UTP-5) as the medium. These may be summarized as follows:

- Signal degradation due to cross-talk (near end and pair-pair).
- Signal attenuation due to lossy medium.
- Limitation of launch power due to FCC compliance.

This paper discusses a method to overcome these problems. The method chosen aims to provide the means to allow a possible low cost but robust implementation using today's available technology, and not requiring advanced computationally expensive DSP implementations.

The method uses a 2 bit to 4 level coding scheme in order to reduce the effective symbol rate on each pair of UTP-5 cable. By using 4 level encoding, the symbol rate on each pair is reduced to either 250MBaud or 125MBaud, depending on the number of pairs. By limiting the transmit launch power of these symbols, and by specifying a maximum cable distance of either 50m or 100m, the power density spectrum can be controlled to meet FCC-B emission standards, while still providing sufficient signal at the receiver to enable reception with a BER of better than 10^{-10} .

This paper first discusses an overview of the technique, followed by a mathematical analysis and associated simulation results, and concludes with empirical results derived from lab experiments using a discrete implementation of the method.

2. 1000BASE-TX Functional Overview

The components required to implement a 1Gbps PHY may be divided into the following sub layers within the PHY layer:

- Gigabit Media Independent Interface (GMII)
- Physical Coding Sublayer (PCS)
- Physical Media Attachment (PMA)
- Physical Media Dependent (PMD)

These components may be mapped into an identical protocol stack as illustrated in IEEE802.3u (100BASE-X), with the modification of the GMII and naturally the MDI. This is illustrated in Fig. 2 below:

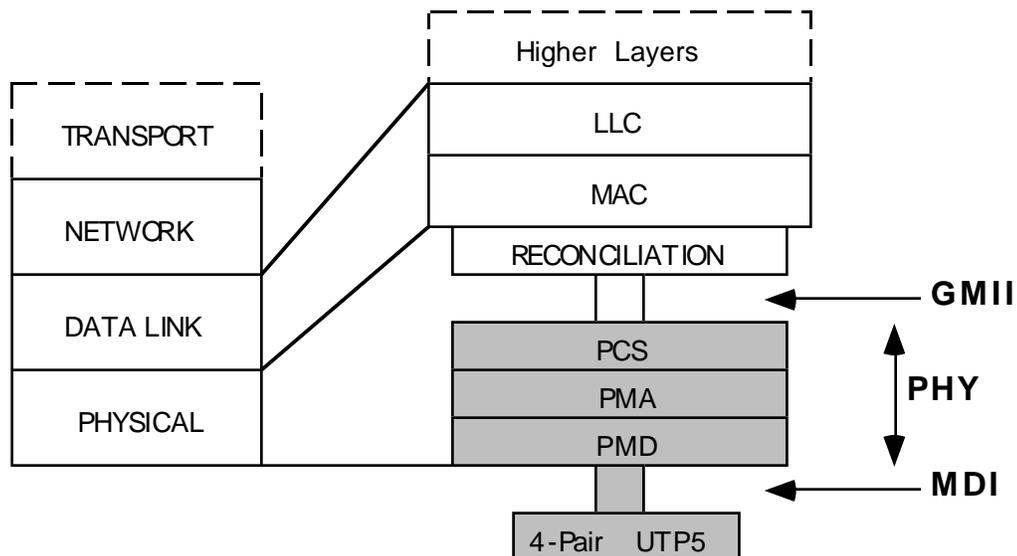


Fig. 2 Architectural Positioning of 1000BASE-TX

The architectural positioning of 1000BASE-TX shown in Fig. 2 maps to the normal view of the protocol stack as described in IEEE802.3 documents. To clarify the proposed solution, the following provides a brief overview without separating the different functions into the appropriate sub-layers.

A simplified block diagram of the 4-pair solution is illustrated in Fig. 3. The 8-pair solution is similar, but uses 4 processing paths for both transmit and receive.

For the sake of brevity, the remainder of this proposal refers to the single cable (4-pair) case, unless otherwise noted.

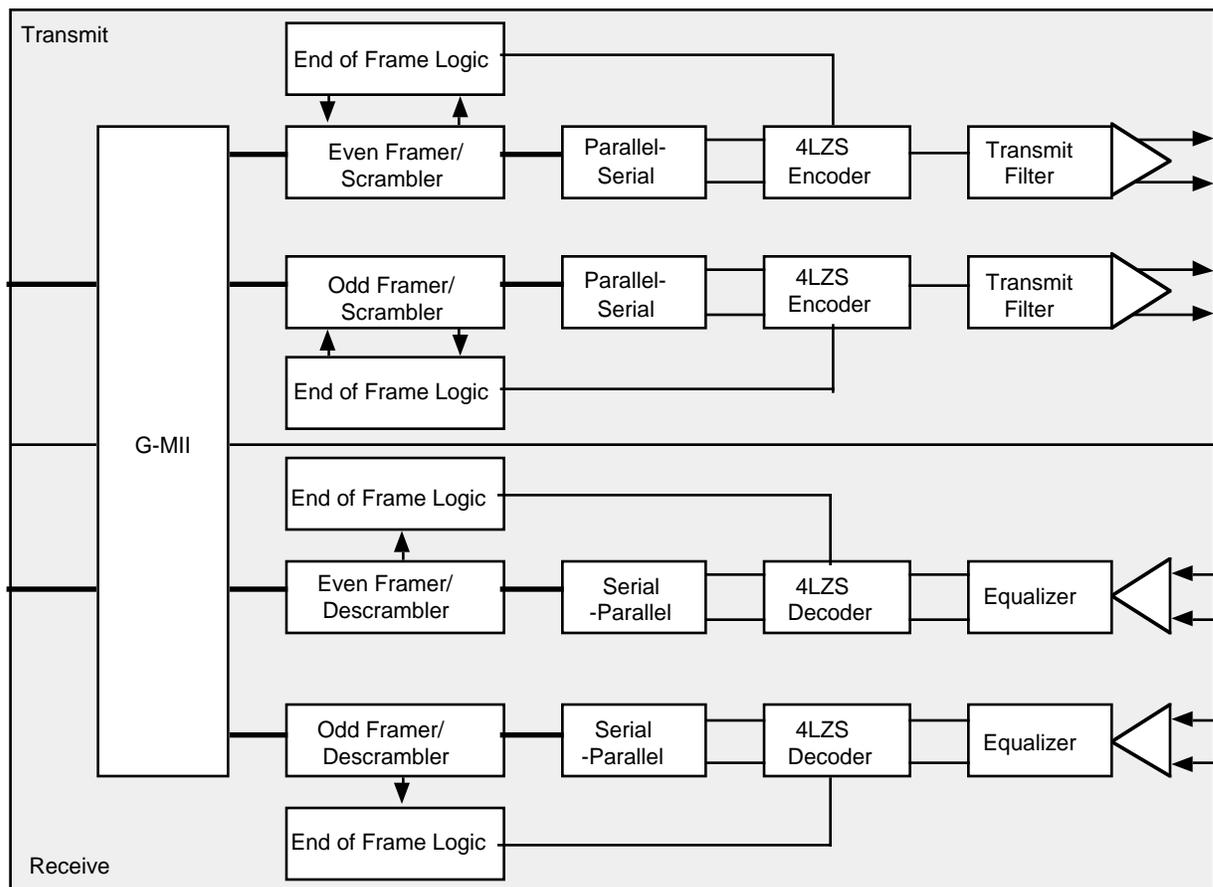


Fig. 3 Block Diagram of 1000BASE-TX

In the transmit path:

- Receive byte aligned data over the GMII at 125Mbytes/s.
- Demultiplex into two parallel byte streams at 62.5Mbytes/s.
- For each parallel stream, map the split ethernet frame into a physical layer stream, by inserting "escape sequences" representing start/end of packet, carrier extension etc as appropriate.
- Scramble the entire frame excluding the escape sequences using a frame synchronous stream cypher scrambler to provide spectral shaping.
- Map the resulting serial bit stream into quats (using 4LZS signal coding), and perform any required quat run length substitution using defined escape sequences.
- Transmit the 4 level signals (from the 2 parallel streams) to the 2 transmit pairs of UTP-5.
- At the end of the packet, transmit an escape sequence representing either end of packet or start of carrier extension, followed by either extension codes or idle codes.

In the receive path:

- For each stream, receive the 4 level signals from the 2 transmit pairs of UTP-5.
- Map the resulting quats into a serial bit stream (using 4LZS signal coding), and perform any required quat run length substitution from defined escape sequences..
- Search for the start of packet by verifying a start of frame escape sequence.
- Descramble the entire frame excluding the escape sequences using a frame synchronous stream cypher descrambler .
- For each stream, search for the end of the packet by detecting an end of packet escape sequence.
- Having received packet and byte alignment, synchronize with the other stream framing block, and reconstruct the original packet (one stream known to contain the first byte).
- Transmit byte aligned data over the GMII at 125Mbytes/s.

3. Existing Ethernet Signal Coding

3.1 4B5B and 8B10B Coding

100BASE-TX uses 4B5B coding in order to convey extra signaling information. Similarly, 1000Base-FX or 1000Base-CX uses 8B10B coding. Both 4B5B and 8B10B coding implies an overhead of 25% of the MII data rate when transmitted on the media via the MDI. These additional codepoints allow the following functions to be carried :

- Data
- Start/End of Stream Delimiter (SSD/ESD)
- Error Conditions
- Idle line
- Provides DC balance

The SSD is used by the receiver to lock onto the received frame, and has the property that allows itself to be uniquely identified in a serial data stream. Once the receiver is in frame sync, the receiver determines the end of packet by searching for the ESD (unique pattern once in sync). The 25% increase in the baud rate with 4B5B or 8B10B makes it increasingly difficult for use at 1000Mbps rates over UTP5 cable.

Taking the definitions of the signaling codes defined in the current fibre or coax based proposals for gigabit ethernet, the following 8B10B signaling codes are defined and conveyed over the channel:

PCS Code	Function	8B10B Encoding	Bytes Required
F	LINK_NOT_AVAILABLE	K28.5 D21.5	2
C	LINK_CONFIGURATION	K28.5 D10.5 config_reg	4
I1	Idle / flip disparity	K28.5 D5.6	2
I2	Idle / same disparity	K28.5 D16.2	2
S	Start of Packet	K27.7	1
T	End of Packet	K29.7	1
R	Carrier Extension	K23.7	1
H	Invalid Code	K30.7	1

Table 1 Control Code Mapping for 8B10B

3.2 Serial/Parallel Conversion (SERDES)

The current fibre proposals assume the use of existing serializers/deserializers (SERDES) components that have been designed for Fibre Channel. These devices assume the 10-bit IDLE words (from 8B10B encoding) are aligned on 20-bit boundaries. As a consequence of this, certain restrictions are applicable in the PCS layer. For example, all transmitted packets start on even numbered characters, and all packets end on odd numbered characters, even if the packet contains an odd number of bytes. In this case, the packet is extended by using the R-codes.

Possible combinations of 10-bit PCS codes are illustrated in Fig. 4 below:

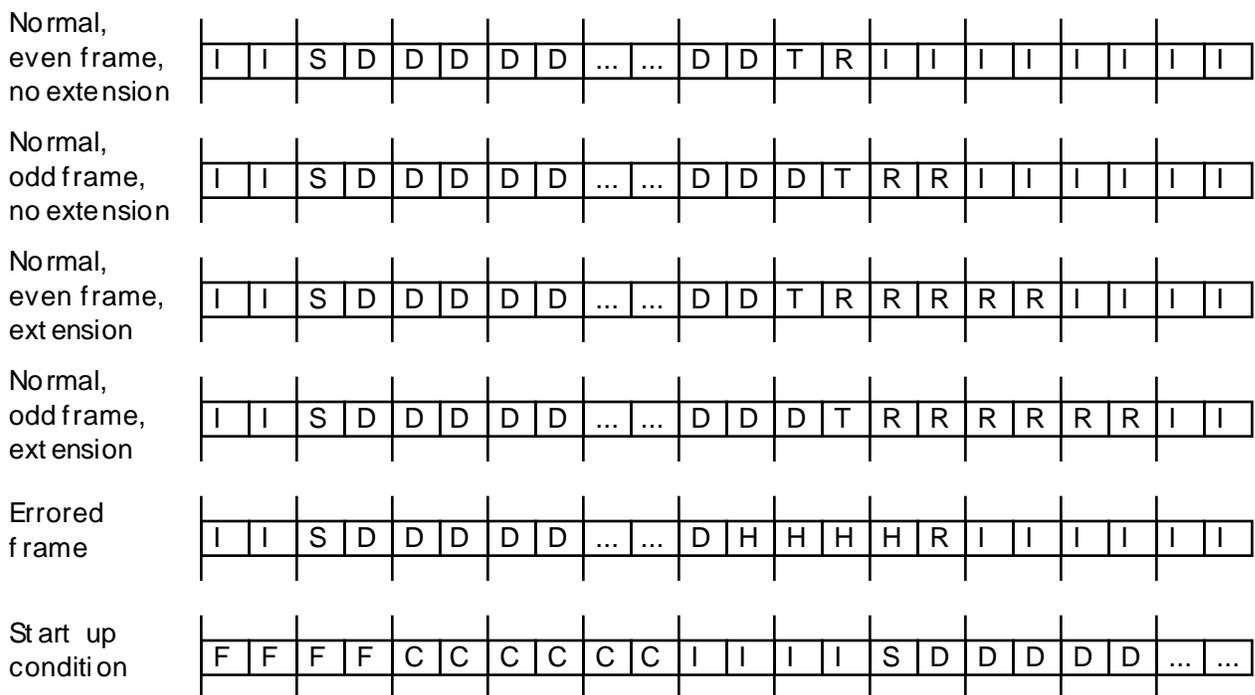


Fig. 4 PCS Mappings for 8B10B codes

4. Signal Coding with 4LZS for 1000Base-TX

4.1 Escape Sequence Overview

The proposed coding scheme allows for a 2x compression of the line rate. 2B1Q line coding maps 2 bits into one of 4 possible levels. The addition of a zero state allows the definition of an escape sequence. The mapping of the bits into symbols, and subsequent voltage levels are illustrated below:

Bits	Symbol	Level (mv)
10	+3	450
11	+1	150
NULL	0	0
01	-1	-150
00	-3	-450

Table 2 4LZS Symbol Code Mapping

The ESC denotes the start of an escape sequence. The escape sequence consists of the ESC plus four additional quats which is equivalent of 1 byte. Escape sequences are denoted as "N+x". This is illustrated in Fig. 5 below:

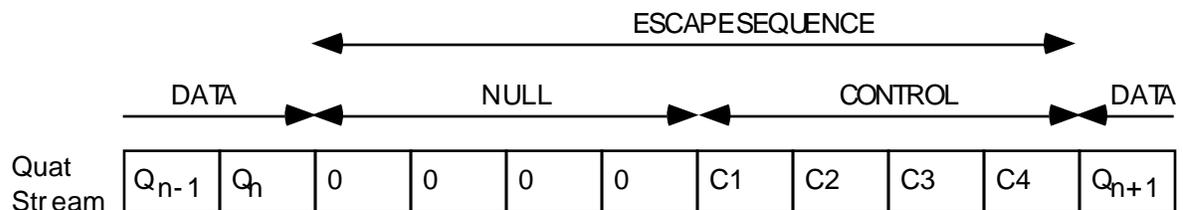


Fig. 5 Escape Code Sequence

Escape sequences are used to convey the following, though not all of the PCS codes require an escape sequence.

- S - Start of Packet
- T - End of Packet
- R - Carrier Extension
- H - Error codes
- F - Link Not Available
- C - Link Configuration
- I - Idle

4.2 Pattern Substitution Escape Sequences

Scrambling is used to ensure a reasonable amount of data randomization, and hence prevent long fixed pattern sequences which would result in long intervals of fixed quats being transmitted. This is undesirable due to the lack of transitions for clock recovery circuits and also because it increases any baseline wander in the receiver. The proposed scrambler uses the same cipher scrambler polynomial a 100BaseTX, namely $x^{11} + x^9 + 1$. The effectiveness of the use of this scrambler is currently being investigated particularly with regard to spectral shaping and density.

In addition to scrambling, escape sequences are used here to reduce this problem. Four binary sequences would cause the transmission of fixed quats. These are all ones, all zeros, an alternating pattern of ones and zeros, or an alternating pattern of zeros and ones. These would map to +1, -3, +3 and -1 respectively.

The transmitter searches for one of the above fixed 16-bit patterns, and on detection, it will substitute the 8-quat word for a 4-quat NULL and an associated 4-quat control code representing one of the substituted patterns. The pattern matching will imply an inherent 16-bit delay in both the transmitter and the receiver. Note that this is not actually a PCS layer function.

4.3 Escape Sequence and PCS Layer Streams

As mentioned previously, all data received at the PCS layer from the MAC will be scrambled prior to being transmitted onto the physical media. Escape sequences are added after scrambling, overwriting the previously scrambled bits. These bits will be regenerated later in the receiver at the far end.

The illustration in Fig. 6 shows how the PCS codes actually map to a PCS stream. Note that the escape sequence is used to define a transition from one "state" to another. For example, an "N+P" escape sequence is used to indicate start of packet from an idle state. The remaining packet bytes are considered normal data. End of packet is indicated by a transition from the packet "P" state to either the idle "I" state or the carrier extension "R" state. Any other escape sequence would be invalid. At any time, pattern substitution may apply using an "N+Sxx" escape sequence. Also note that all escape sequences are 2 bytes in length

Data following an N+C escape sequence will consist of a 16-bit "configuration register". This 4 byte configuration sequence will repeat until acknowledged. The data after an N+P will consist of packet data.

Fig. 6 below illustrates 2 example streams. The first stream contains a packet with no carrier extension, and the second stream has carrier extension.

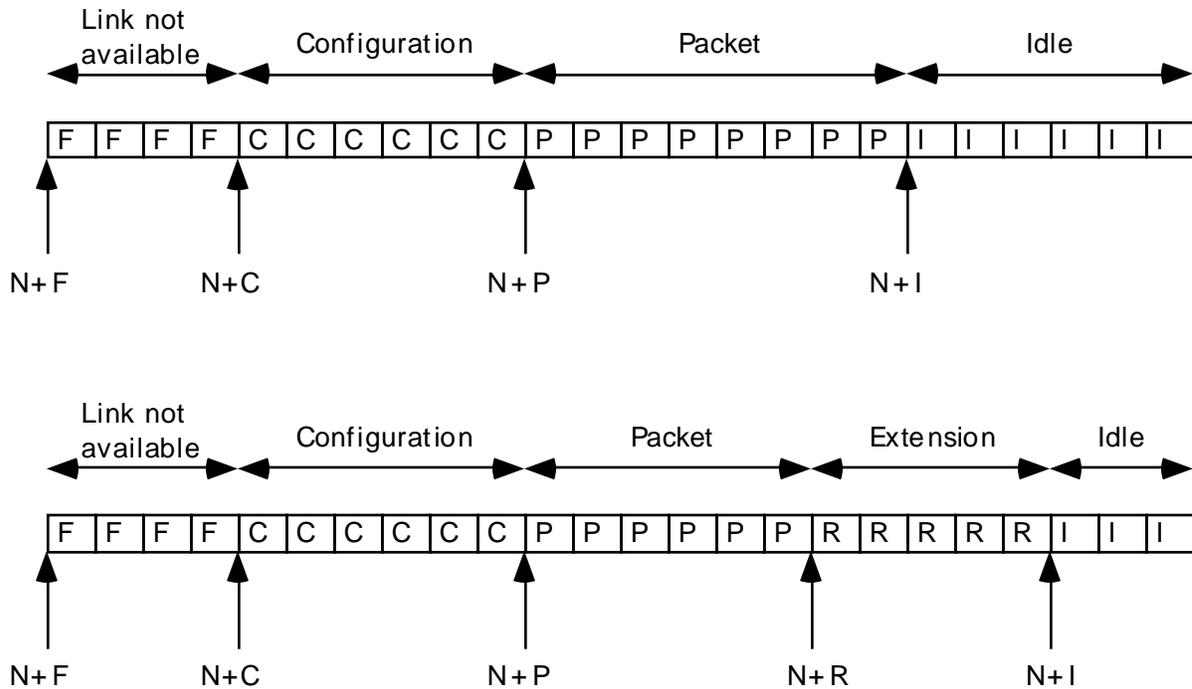


Fig. 6 Example 4LZS PCS Layer Streams

4.4 Escape Sequence Mappings

As described earlier, each escape sequence consists of a 4 quat NULL followed by a 4 quat control byte. This control byte has 256 possible codings. One requirement which these 4LZS control codes have an impact on is that of "Hamming Distance". In relation to gigabit ethernet, this requirement is restated verbatim as follows: "A minimum of four bit cells in error shall be necessary for an undetected error to occur (Hamming distance 4).".

Data sequences that transform onto other data sequences do not pose any concern here as these errors will be detected by the CRC-32 FCS. Data sequences that transform to control codes however, can "restructure" an Ethernet frame in the case of end-of-packet or carrier extension. Control codes are not protected by the FCS; some error protection will be needed to detect errors that transform one control code to another. In some cases, erroneous control codes can be detected from the context in which they appear. For example, it does not make sense to have carrier extension prior to start-of-packet. In other cases however, control code errors may pass through undetected or not be detectable in a timely manner (e.g., carrier extension -> end-of-packet). Errors that transform a control code to a data sequence are also of concern in the case of carrier extension and end-of-packet. These are at least some scenarios in which the Hamming distance objective may be compromised.

Details on selection of the control codes is provided in reference [7]. In summary, the overall approach has been to make all control codes sufficiently distinct from data sequences, and then to make the control codes sufficiently distinct from each other. This is accomplished in the first case by zero-state escape sequences, and in the second case by using an (8, 4) extended Hamming code.

The following table illustrates the mappings for the escape sequences defined not only for the PCS layer, but also for the pattern substitution. Note that PCS codes which are escaped using a NULL are denoted as "N+x", where x is the "state" identifier.

PCS Code	Function	4LZS Control Byte Encoding (NULL,C1,C2,C3,C4)							
N+F	Start of "F"	0	0	0	0	-3	-3	-3	-3
F	Link Not Available	-3	-3	-3	-3	-3	-3	-3	-3
N+C	Start of "C"	0	0	0	0	+1	-1	-3	-1
C	Link Configuration	"16-bit Configuration Register"							
N+I	Start of idle (EOP, no ext)	0	0	0	0	-1	+1	-3	+3
I	Idle	-1	+1	-3	+3	-1	+1	-3	+3
N+P	Start of "P" (SOP)	0	0	0	0	+3	+3	-3	+1
P	Packet Data	"Packet data"							
N+R	Start of "R" (EOP, with ext)	0	0	0	0	+3	+1	-1	-3
R	Carrier Extension	+3	+1	-1	-3	+3	+1	-1	-3
N+H	Start of "H"	0	0	0	0	-1	+3	-1	-1
H	Invalid Code	-1	+3	-1	-1	-1	+3	-1	-1
N+S11	Substitute "11...."	0	0	0	0	-3	-1	-1	+1
N+S00	Substitute "00...."	0	0	0	0	-3	+1	+3	-1
N+S01	Substitute "01...."	0	0	0	0	+1	+3	+3	-3
N+S10	Substitute "10...."	0	0	0	0	+1	-3	-1	+3

Table 3 4LZS Escape Sequence Mappings

Note that after transmission of an escape sequence that indicates a change of state and a subsequent repeating pattern (i.e. F, I, R, H), then the control byte after the NULL byte will simply be repeated. Either of these repeating patterns may be truncated to a single byte at any time, depending on the alignment of data. In other words, the transmission of data or codes does not need to be aligned on 16-bit boundaries.

The mapping are illustrated graphically in Fig. 7.

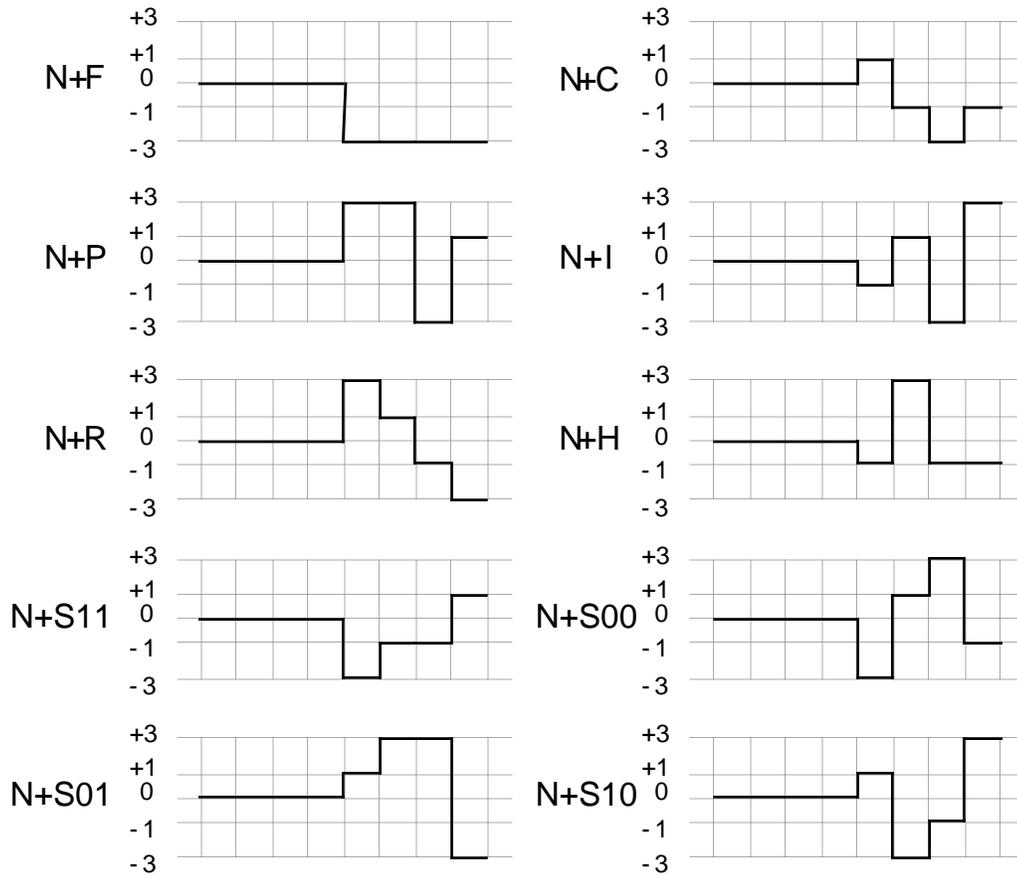


Fig. 7 Quat Domain Representation of the Escape Sequences

4.5 Physical Layer Streams for Single/Dual Cables

An example of physical layer streams for both single and dual cable configurations is illustrated in Fig. 8 and Fig 9 respectively below. In the diagrams note how the preamble bytes are overwritten by start of packet escape sequences. Also note how the inter packet gap (IPG) is used to allow end of packet delineation by a similar escape sequence. Both these diagrams illustrate 2 consecutive packets, separated by the minimum IPG of 12 bytes. They also illustrate the case of the packet being large enough not to require carrier extension. This is shown by the N+I escape sequence. For carrier extension, the escape sequence would be N+R, followed by R-codes to denote carrier extension. This would then be followed by an N+I escape sequence, followed by I-codes until the start of the next packet.

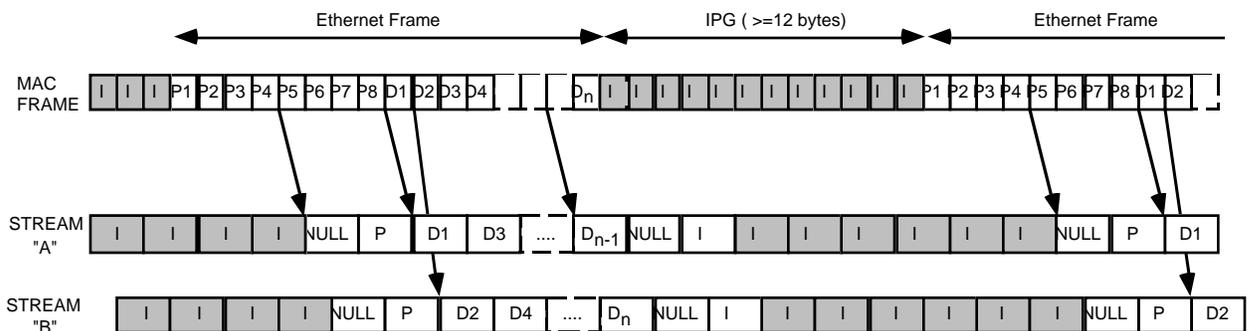


Fig. 8 Single Cable Physical Layer Streams A and B

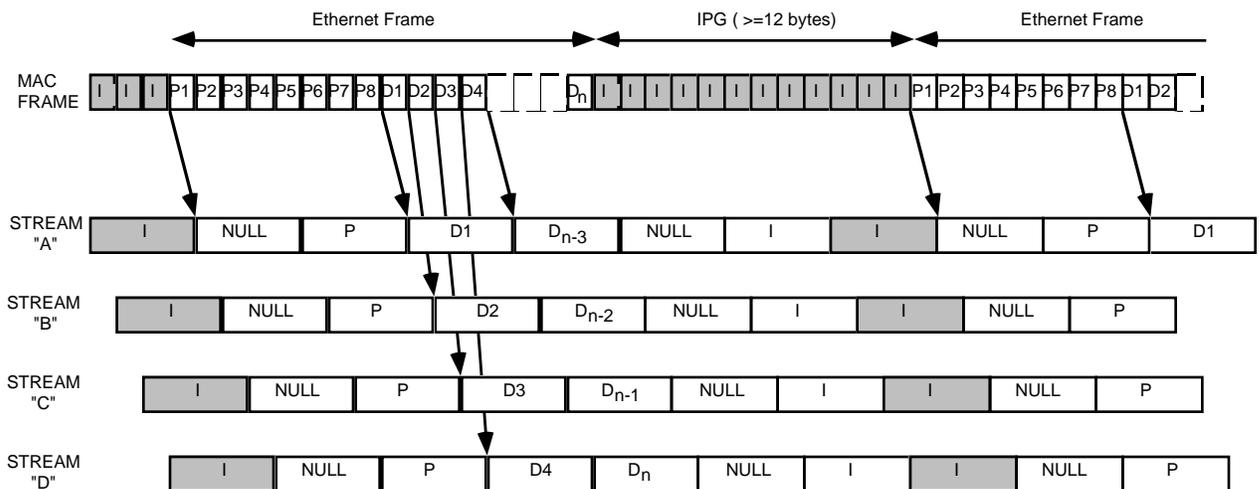


Fig. 9 Dual Cable Physical Layer Streams A, B, C and D

4.6 Frame Scrambling

Scrambling is used to spread the spectrum of transmitted data so as to reduce EMC problems associated with broadband data transmission over unshielded media. Quasi-periodic patterns in the source data stream can introduce emission spikes. By randomizing these patterns through scrambling, the emission spectrum is smoothed out. Scrambling can also improve data transmission characteristics by reducing jitter and inter-symbol interference. Runs (sequences of unchanging bits) are also broken up by scrambling, resulting in a more reliable stream for clock recovery. Scrambling is also used to provide some measure of link security.

Frame synchronous cipher stream scrambling is recommended to avoid error multiplication inherent with self-synchronous scrambling, and possible bandwidth expansion as well as additional implementation complexity resulting from distributed sample scrambling. Frame synchronous scrambling is specified for 100Base-Tx Ethernet in the ANSI X3.263-1995 TP-PMD standard [9]. The generating polynomial is given by $1 + x^9 + x^{11}$. It is recommended to use this same polynomial.

The killer packet problem is addressed by a substitution scheme in which constant runs are replaced by uniquely identifiable patterns to restore the requisite transition density required for clock recovery. Since constant runs also occur, however improbable (as analyzed in [8]) in completely random, scrambled data, the substitution code is also of incidental utility to *guarantee* transition density, a characteristic of 8B10B-coded data transmission.

Excluding killer packet attacks, the scrambled data stream will appear to be random if either the scrambled or the unscrambled data pattern is random. A well-designed scrambler should produce, for all practical purposes, a random bit stream. Likewise, the fields of Ethernet frames are generally random. The principal benefit then from scrambling comes from randomizing repeated patterns, such as the preamble, carrier extension and idle sequences. In the absence of killer packet attacks, it ought to suffice to study only the case where the scrambled stream is random.

Table 4 shows the run length distribution without substitution. While a "run" usually means a constant symbol stream, the runs "seen" for clock recovery purposes may differ however, depending on implementation. A simplified edge detection circuit may only be detecting transitions between pairs of levels (e.g., (+3,-3) and (+1,-1)) for example. Accordingly, two sets of probabilities are shown. The probabilities for $p = 0.25$ correspond to constant symbol runs (applicable to clock recovery circuits that detect all transitions), while those for $p = 0.5$ correspond to runs consisting of any one of two symbol pairs. It can be seen that in either case, long runs are extremely improbable. For the case where substitution is used to bound run lengths, the corresponding distributions have also been evaluated to completely characterize the resulting truncated runs.

A full analysis of this can be found in [8], but in summary substitution on a run of 8 quats (2 bytes) aligned on 16 bit boundaries bounds the maximum run length to 14 bytes.

Run Length	Probability	
	$p = 0.25$	$p = 0.5$
1	0.75	0.5
2	0.1875	0.25
4	0.01172	0.0625
8	4.578e-5	0.00391
16	6.985e-10	1.526e-5
32	1.626e-19	2.33e-10
64	8.816e-39	5.42e-20

Table 4. Run Length Distribution without Substitution.

It can be seen in the above table that the number of long runs (e.g., 32, or 64 quats) becomes statistically a very small fraction of the total number of runs. The corresponding proportion of time occupied by long runs is very small. As an example, it can be shown that the proportion of time occupied by runs longer than 64 quats for $p = 0.25$ is 1.44×10^{-37} , or roughly 1 run longer than 64 quats in 4.4×10^{20} years.

SONET requires that clock recovery be possible for runs of up to 72 bits. PMC-Sierra's chips are typically capable of maintaining sync for runs of even up to 80 bits (line symbols) long. It is clear that in general, if killer packets are excluded, a good scrambler alone (without run substitution) should be sufficient to ensure reliable clock recovery.

4.7 Comparison of 4LZS against NRZ

The spectral shape of a 4LZS line code compared with an NRZ line code is illustrated in Fig. 10 below:

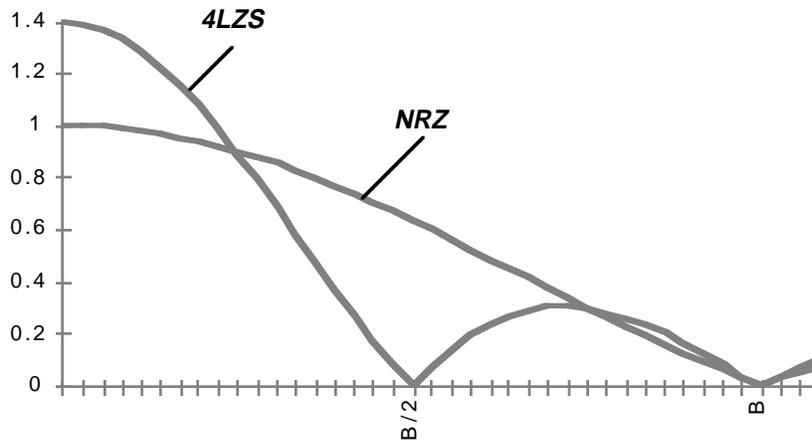


Fig. 10 Spectrums of 4LZS and NRZ Line Codes

The resulting ideal eye diagram of a 4LZS line code is shown in Fig. 11 below.

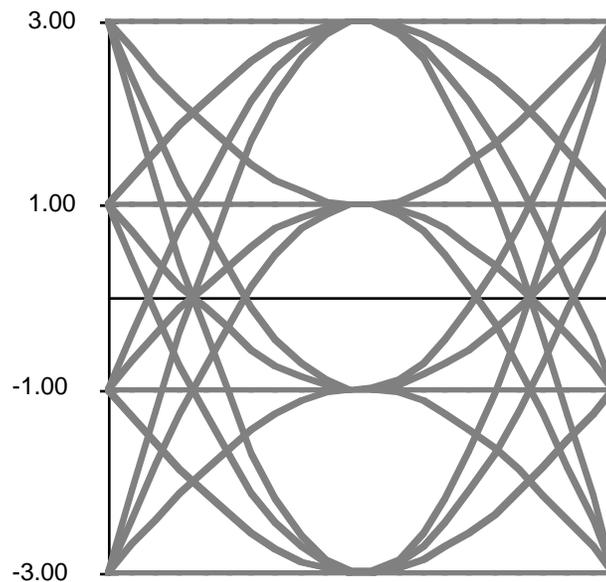


Fig. 11 Ideal 4LZS Eye Diagram

4.7.1 Theoretical Analysis of 2B1Q against NRZ

Additional SNR required at the slicer may be computed as the intrinsic coding power density, which is the "symbol variance":

$$\text{SNR} = \sigma_s^2 / \sigma_n^2 ; \Delta\text{SNR} = \sigma_{s2B1Q}^2 / \sigma_{sNRZ}^2$$

where, $\sigma_s^2 = 1/L_C \sum_i \langle e_i^2 \rangle$

$$\sigma_{s2B1Q}^2 = 1/4 \{ (-3)^2 + (-1)^2 + (1)^2 + (3)^2 \} = 5$$

$$\sigma_{sNRZ}^2 = 1/2 \{ (-1)^2 + (+1)^2 \} = 1$$

Therefore,

$$\Delta\text{SNR} = \sigma_{s2B1Q}^2 / \sigma_{sNRZ}^2 = 10\log(5/1) = \mathbf{7dB}$$

In addition to the SNR penalty at the receiver, noise and loss are proportional to the symbol period and cable geometry.

Flat channel noise power is proportional to the symbol rate (1/T) :

$$\begin{aligned} \Delta\sigma_{nflat} &= -10 \log [T_{2B1Q} / T_{NRZ}] \\ &= \mathbf{2 dB} @ 250\text{Mbaud vs } 155\text{Mbit NRZ} \end{aligned}$$

Self NEXT noise power is proportional to the cubed square root of the symbol rate ($T^{3/2}$) :

$$\begin{aligned} \Delta\sigma_{nsnext} &= -15 \log [T_{2B1Q} / T_{NRZ}] \\ &= \mathbf{3 dB} @ 250\text{Mbaud vs } 155\text{Mbit NRZ} \end{aligned}$$

NEXT noise power is frequency and cable geometry dependent :

$$\begin{aligned} N(f) &= \chi f^{3/2} G(f) ; \chi = 6.31 \times 10^{-7} \text{ for UTP5} \\ &= \mathbf{3 dB} @ 250\text{Mbaud vs } 155\text{Mbit NRZ} \end{aligned}$$

In summary, 250Mbaud 2B1Q penalties over 155NRZ are:

- 6dB attenuation,
- 2dB flat channel noise,
- 3dB self NEXT,
- 3dB pair-pair NEXT
- 7dB at the slicer

Totaling a **21dB** penalty. But how far is 21dB?

Characterizing power spectral density at the centre frequency:

$$\text{NRZ}(155) \sim 1.967 \sqrt{f} + 0.023 f + 0.05 / \sqrt{f} = 19.1 \text{ dB}/100\text{m} @ 77.5 \text{ MHz}$$

For 2B1Q (dB/100m) x D metres = NRZ (dB @ 100m)

$$D = 19.1 / (19.1 + 21) = 48 \text{ metres}$$

Therefore, theoretical performance:

$$\mathbf{100\text{m of 155 NRZ} = 48\text{m for 250Mbaud/pair 2B1Q}$$

4.8 Receiver Line Equalization

At the receiver, line equalization will be required in order to compensate for the line characteristics. This may be done using either adaptive or fixed equalization. The exact method chosen will be implementation specific and is not discussed further here.

4.9 Single or Dual Cable Operation

As mentioned earlier in this paper, an increase in the distance (or even speed) objectives is possible using dual cables having a total of 8-pairs of UTP-5.

In this scenario, each cable (of 4-pairs) would equally share the bandwidth, both cables having 2 transmit and 2 receive pairs.

An illustration of the distance/rate trade-off for both single and dual cable cases is illustrated in Fig. 12 below.

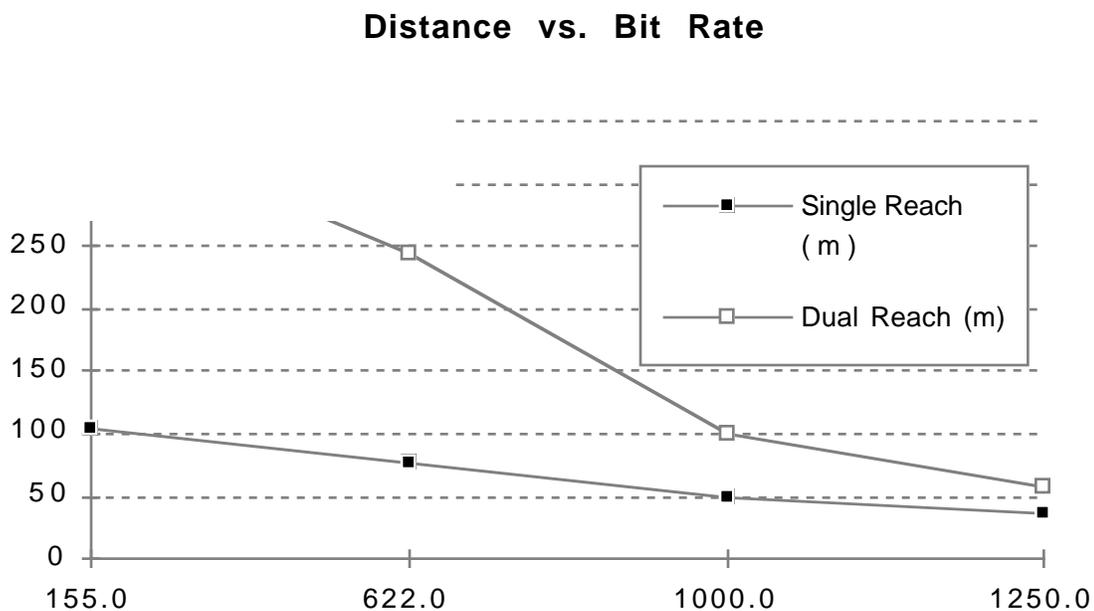


Fig. 12 Illustration of Speed vs Distance for Single/Dual 4-pair Cables

5. Gigabit Media Independent Interface (GMII)

The function of the GMII is to interface the PHY layer device to a higher layer function (MAC layer, via a reconciliation layer). The GMII defined here is a simple extension to the existing MII as defined in IEEE802.3u for 100BASE-X. The only significant change is the requirement to transfer 1Gbps full duplex over this interface, which using the existing nibble based interface would clearly not be practical as it would require a 250MHz clock.

The proposal here is to extend the nibble interface to a byte wide interface. This would allow 1Gbps full-duplex support using a 125MHz clock, which simplifies the required interface and allows implementation using relatively easily available technology. An additional clock is also required for retiming of the transmit data signals. This is described later.

Key features of this interface are :

- Capable of supporting 10, 100, or 1000 Mbps data rates.
- Data and delimiters synchronous to clock.
- Independent 8-bit wide transmit and receive data paths (drops back to nibble based interface for 10 and 100 Mbps).
- Uses TTL signal levels compatible with common digital ASIC CMOS processes.
- Provides simple management interface.

5.1 GMII Functional Overview

Information is conveyed between the MAC and PCS layers via a Gigabit Media Independent Interface (GMII). This information will consist of either data or control. How the control information is conveyed over the GMII is currently under debate, but the current draft proposal is outlined below. This interface uses reserved signal combinations for extensions, to allow signaling information to be carried over the GMII.

The exact method of indicating the various signaling information elements is not too critical here. However, the method for transmitting the information on the line is. It is assumed that this information will be passed to the PCS layer using some robust mechanism.

The control mapping for both transmit and receive paths currently defined is illustrated in Table 5 and Table 6.

TX_EN	TX_ER	TXD[7:0]	Meaning
0	0	00 - FF	Normal inter-frame
0	1	00 - 0E	Reserved
0	1	0F	Carrier extension
0	1	10 - FF	Reserved
1	1	00 - FF	End Frame Delimiter
1	1	00 - FF	Carrier Extension

Table 5 Control Mapping over Transmit GMII

RX_EN	RX_DAV	RXD[7:0]	Meaning
0	0	00 - FF	Normal inter-frame
0	1	00	Normal inter-frame
0	1	00 - 0D	Reserved
0	1	0E	False carrier indication
0	1	0F	Carrier extension
0	1	10 - FF	Reserved
1	0	00 - FF	Normal data transmission
1	1	00 - FF	Data reception with errors

Table 6 Control Mapping over Receive GMII

5.2 GMII Signals and Timing Requirements

The following signals are required to implement the GMII. The main enhancements are signals to extend the transmit and receive nibble based interface to a byte wide interfaces (TXD and RXD). An additional output clock (TX_CLK) is also provided by the MAC which would be retimed to the TXD signals to reduce any output jitter. This clock would be regenerated from a reference provided by the PHY layer (TX_CLKR). The complete GMII signals are:

TX_ER	RX_ER	CRS
TX_EN	RX_DV	COL
TX_CLK	RX_CLK	MDC
TXD[7:0]	RXD[7:0]	MDIO
TX_CLKR		

In common with the MII, TX_ER, TX_EN and TXD are synchronous to TX_CLKR (TX_CLK in the MII). Similarly, RX_ER, RX_DV and RXD are synchronous to RX_CLK. CRS and COL are asynchronous to all clocks. It is proposed that the flow through clocks are defined at the GMII. That is, the PHY device would generate the RX_CLK, whereas the MAC layer device would generate the TX_CLK, but this would in turn be generated from the reference clock TX_CLKR provided by the PHY.

The concept of flow through timing is illustrated in Fig. 13 below:

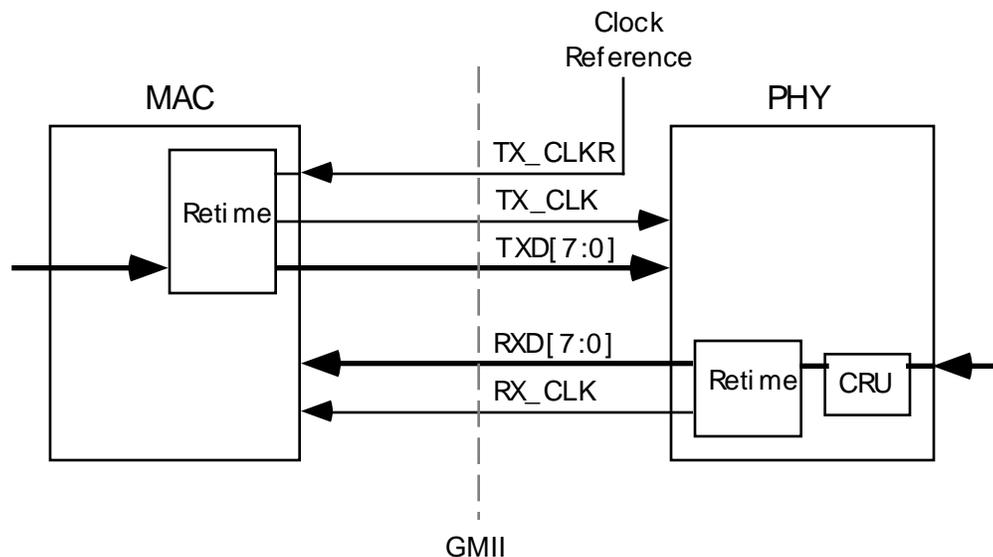


Fig. 13 Flow Through Timing at the GMII

5.3 GMII Functional Timing

The functional timing of the GMII is identical to that for the existing MII for 100BASE-X (TX_CLKR would be the same as TX_CLK). This is shown for both transmit and receive paths in Fig. 14 and Fig. 15 below. Note that CRS and COL are shown simply to indicate the asynchronous nature of the signals and should not necessarily be read in conjunction with the other waveforms :

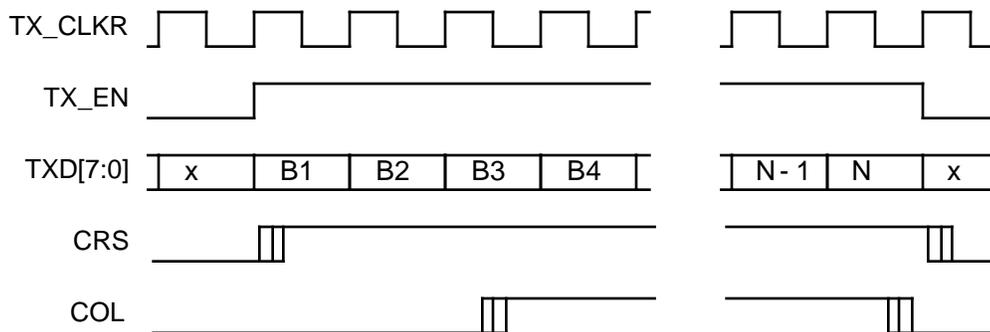


Fig. 14 GMII Transmit Timing

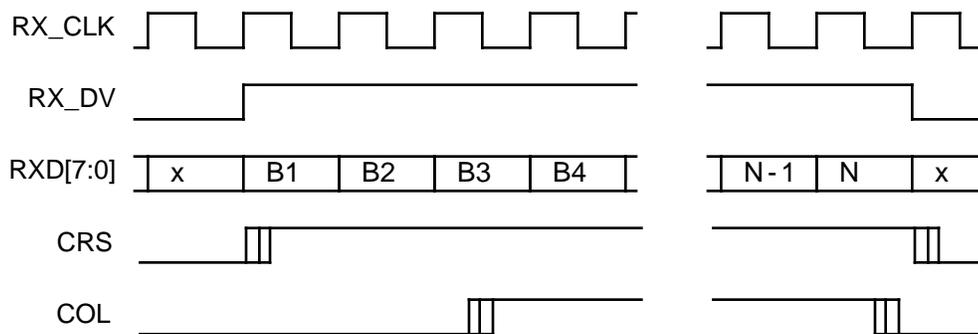


Fig. 15 GMII Receive Timing

5.4 GMII AC Timing

The AC timing and drive requirements for the GMII will be different from MII. With a receive and transmit clock frequency of 125MHz, the period is 8ns. Suitable propagation (T_{PD}), setup (T_{HD}) and hold (T_{SU}) times for synchronous signals across this interface are defined below:

Synchronous to TX_CLKR: (TX_ER, TX_EN, TXD[7:0]) $T_{PD} = \text{TBD}$
 $T_{HD} = \text{TBD}$
 $T_{SU} = \text{TBD}$

Synchronous to RX_CLK: (RX_ER, RX_DV, RXD[7:0]) $T_{PD} = \text{TBD}$
 $T_{HD} = \text{TBD}$
 $T_{SU} = \text{TBD}$

This is illustrated in Fig 16 below:

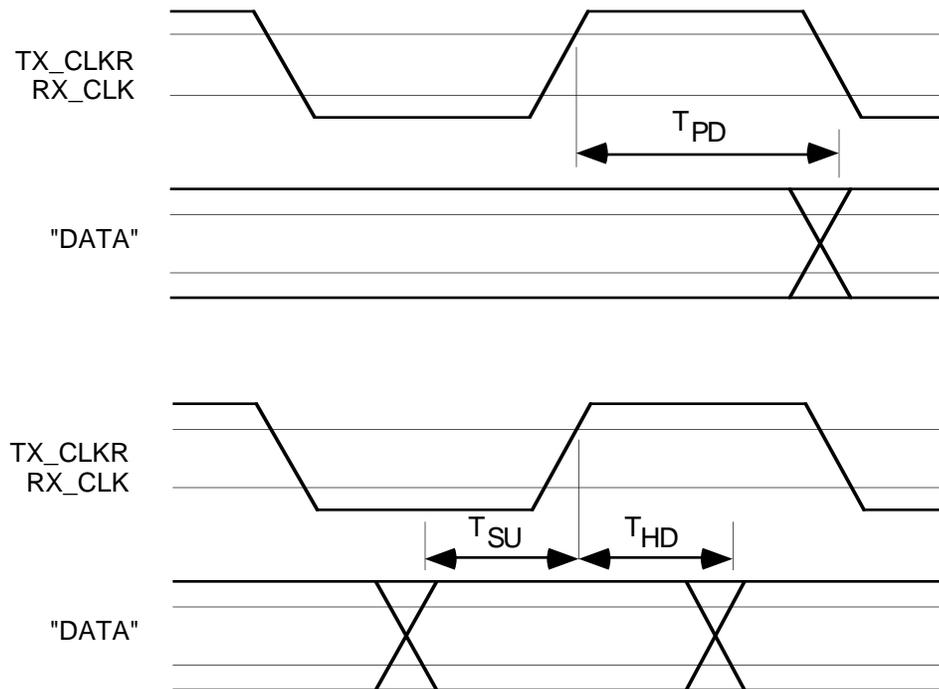


Fig. 16 GMII AC Timing Characteristics

5.5 GMII Management Interface

MDC and MDIO are used to provide a management interface. Control register 0 and Status register 1 are referred to as the "basic register set". Bits would need to be defined in both of these registers to control/indicate 1Gbps support. The functionality of the interface is unchanged except for the addition of these registers.

6. Simulation Model and Results

6.1 Overview

Characterization of UTP-5 cable was performed from (DC) to 300 MHz. Using a step input, several different samples of both Belden "DataTwist-5" and AT&T "Systemtwist" 1061B were measured at lengths of 0, 20, 40, 60, 80, 100 and 120 meters. Polynomial curve fitting of the measured data was performed using the MATLAB "Polyfit" function. The worst case error of the polynomial against measured data was less than 0.1%. Using MATLAB, the step response polynomials were transformed to the frequency domain in order to obtain the UTP-5 cable transfer function (magnitude and phase vs frequency) for each cable length.

Two 4-level stimulus functions were generated assuming a 30% UI (1.2 ns in a 4.0 ns period) edge interval. The two different stimulus functions cover all transitions and resulting trajectories and are DC balanced. No allowance for baseline wander was included. The time domain stimulus functions were transformed to the frequency domain using the MATLAB "FFT" function.

Using the cable transfer function for various cable lengths and the 4-level frequency domain stimulus, the cable response as a function of length was generated. This result was transformed back to the time domain. (The MATLAB inverse function was used.) Eye diagrams were generated by superimposing each of the transitions of the time domain results.

An unequalized eye generated from a simulated 20m cable is illustrated in Fig. 17.

The simulation results were promising enough to believe that a physical prototype would be worth investigating. This work is described next.

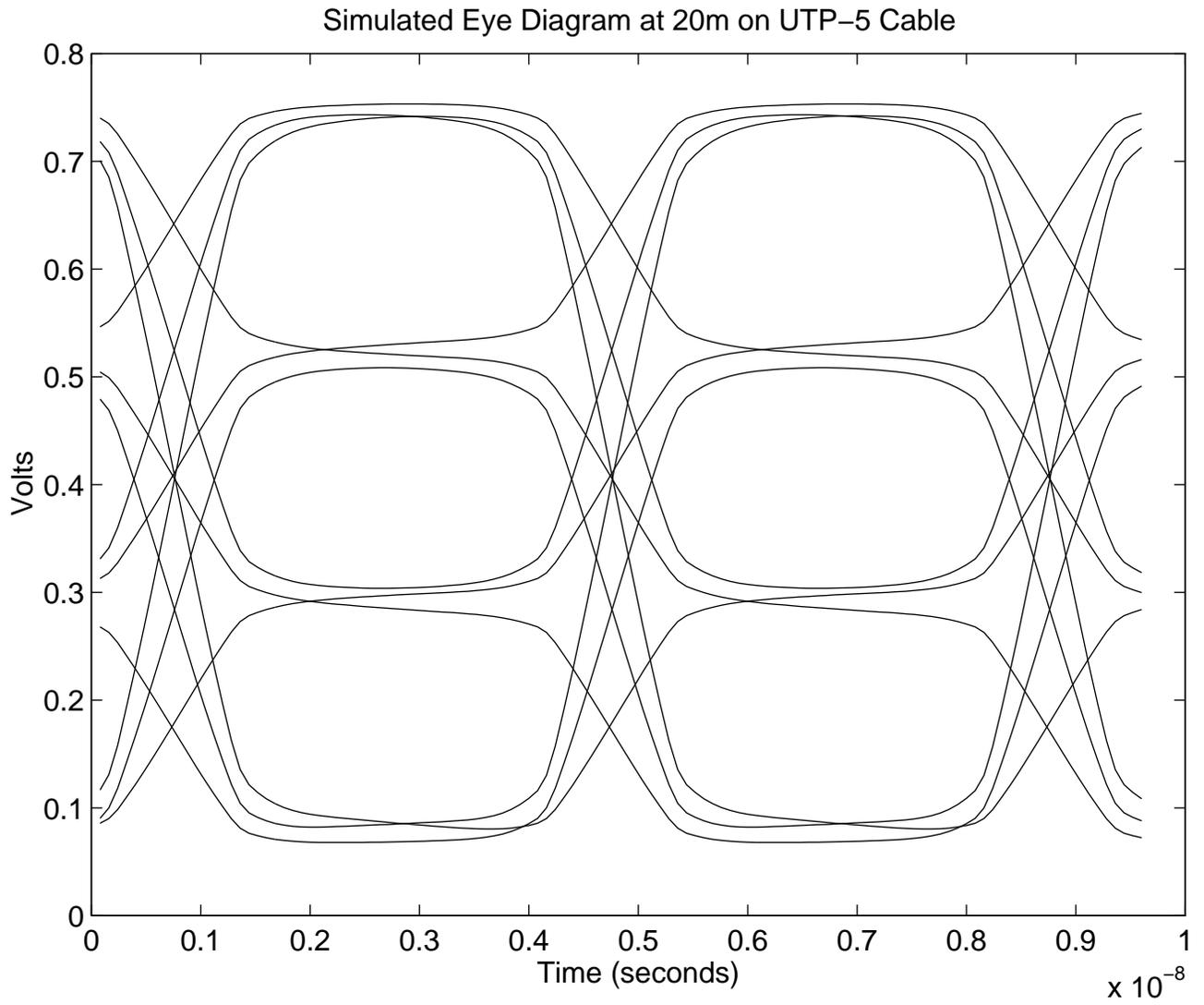


Fig. 17 Simulated Eye Diagram at 20m on UTP-5 Cable

7. Hardware Evaluation and Results

7.1 Overview

A discrete implementation of the PMD discussed in this proposal was built to analyze the performance and behavior of the proposed method, and observe the eye opening under realistic conditions.

A simplified block diagram of the evaluation board is shown in Fig. 18 below:

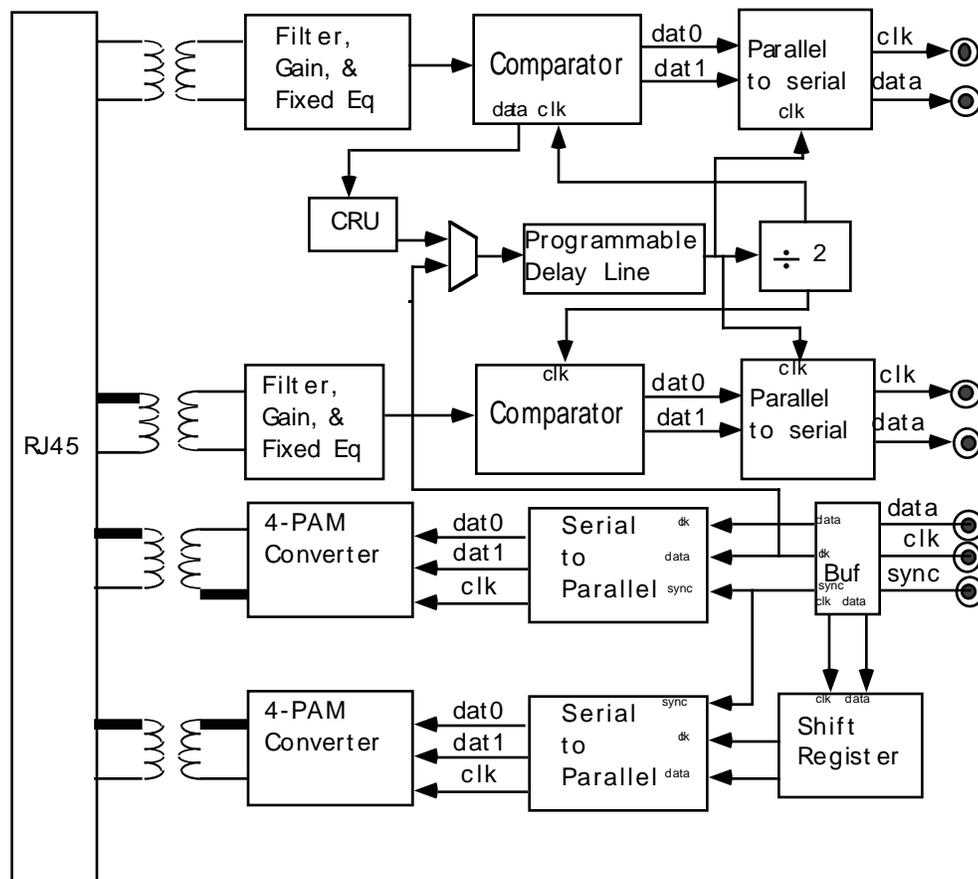


Fig. 18 PMD Evaluation Board Block Diagram

Features of the eval board are:

- 2 transmit and 2 receive channels for NEXT measurements (self, pair-pair, far)
- Each channel operates at 500+ Mbps
- Discrete fixed equalizer

7.2 Test Configuration

Various tests were performed while transmitting a 1Gbps serial data stream from a BER Tester and observing the received eye diagram over various cable lengths. Fixed equalization was used at the receiver on tests T3 and T6 only. The other tests had no implicit equalization.

Three different test configurations were used.

Configuration 1 uses only the transmit section of the evaluation board. The eyes are observed after some defined length of cable, plus 2 x 3m patchcords. Test configuration 1 is illustrated in Fig. 19.

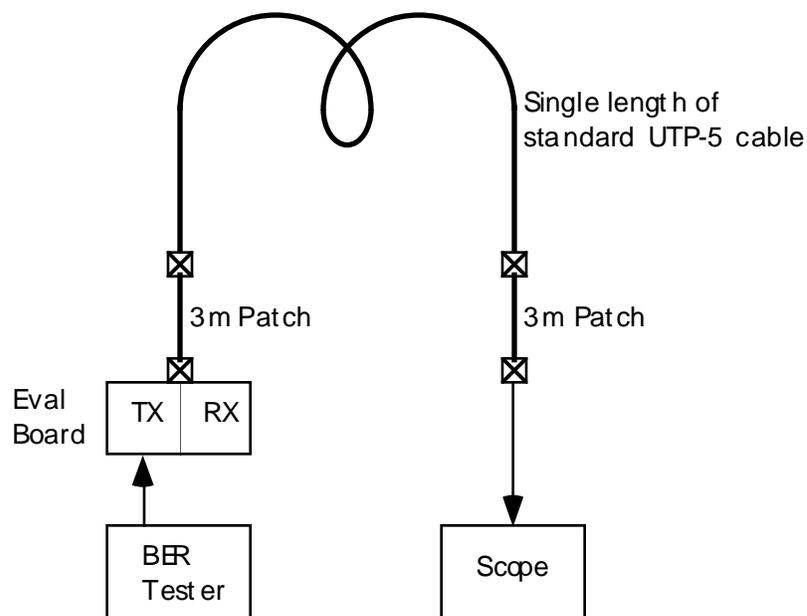


Fig. 19 Test Configuration 1

Configuration 2 uses both the transmit and receive sections of the evaluation board. The eyes are observed after some defined length of cable, plus 2 x 3m patchcords. In addition, the eyes are observed at the output of the receive pre-amplifier. Test configuration 2 is illustrated in Fig. 20.

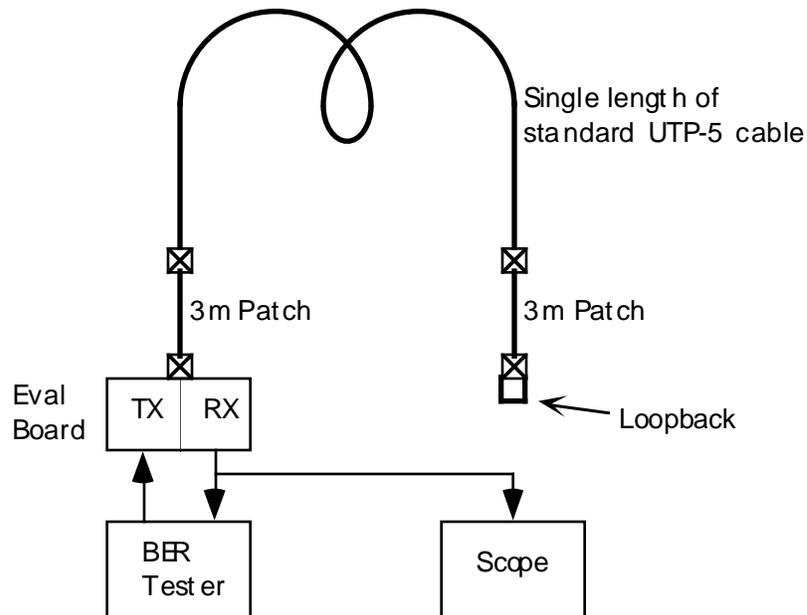


Fig. 20 Test Configuration 2

Configuration 3 provides a more typical setup and avoids loss in the loopback connection used in configuration 2. It also uses the transmit and receive sections of two independent evaluation boards. The eyes are observed after some defined length of cable, plus 2 x 3m patchcords. In addition, the eyes are observed at the output of the receive pre-amplifier. Test configuration 3 is illustrated in Fig. 21.

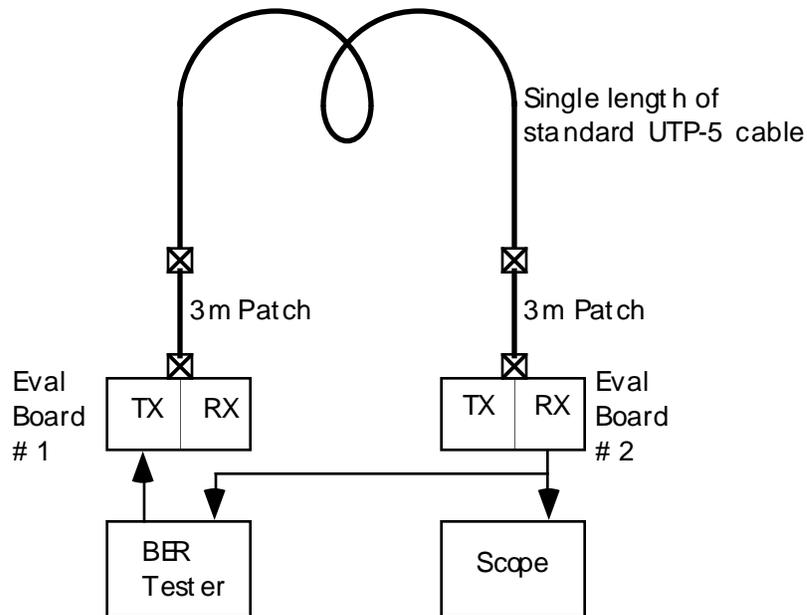


Fig. 21 Test Configuration 3

7.3 Test Procedure

Tests were run using varying cable lengths from 2 major suppliers of UTP-5 cable (AT&T Systimax 1061B, and Belden Datatwist-5 1583A).

The BER tester was used to provide different data patterns (fixed, PRBS-7, PRBS-15 and PRBS-23).

A subset of the individual tests are illustrated in the following table. All of these tests used configuration 3.

Test Num	Test Pattern	Fixed Line Equalization	Horizontal Distance	Observation Point	Eye Plot
T1	PRBS-7	No	50	After Transmit Transformer	Plot 1
T2	PRBS-7	No	50	After Receive Transformer	Plot 2
T3	PRBS-7	Yes	50	Receive amp output	Plot 3
T4	PRBS-23	No	50	After Transmit Transformer	Plot 4
T5	PRBS-23	No	50	After Receive Transformer	Plot 5
T6	PRBS-23	Yes	50	Receive amp output	Plot 6

Table 7 Test Matrix for Distance Evaluation

7.4 Test Results

The resulting eye diagrams for the test described above are shown in Plots 1 - 6.

Plots 1 & 4 are observed at the output of the transmit transformer, immediately before the cable, using a PRBS-7 and PRBS-23 sequence respectively.

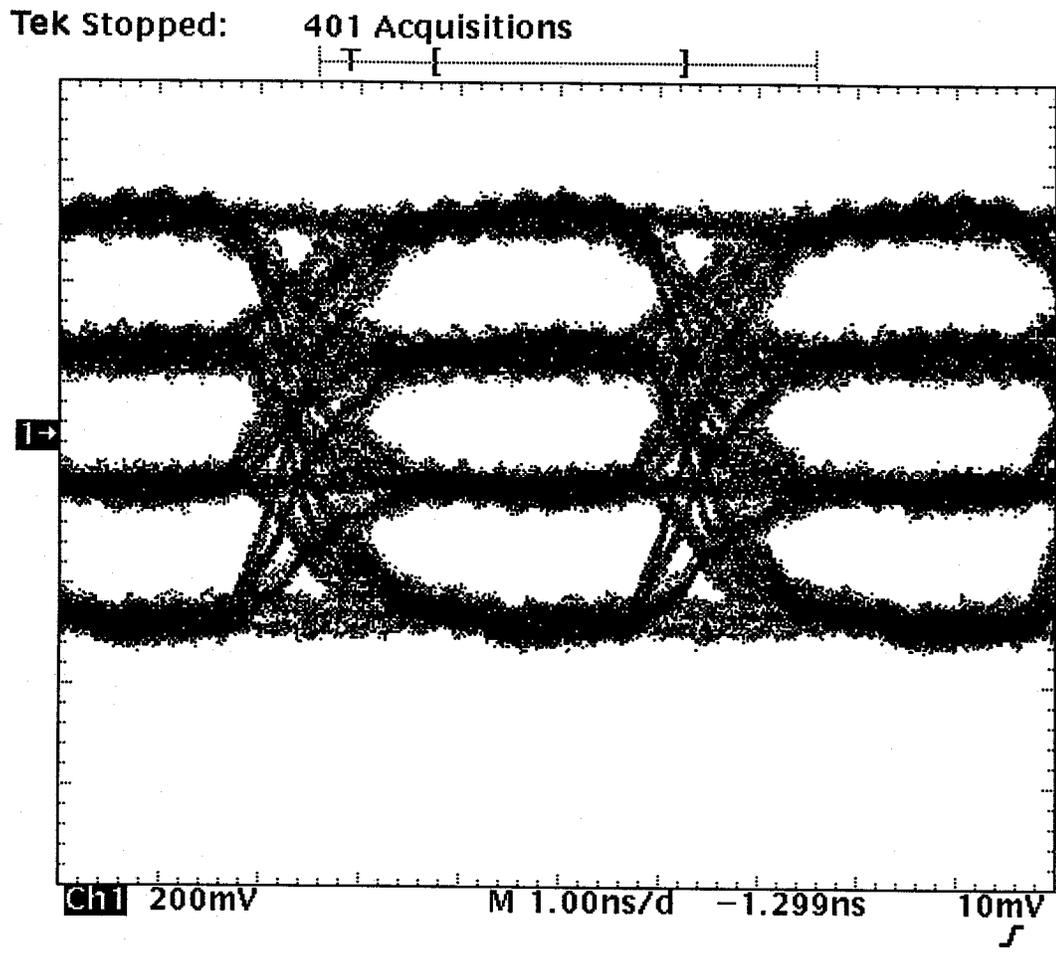
Plots 2 & 5 are observed at the output of the receive transformer, prior to any amplification or equalization, using a PRBS-7 and PRBS-23 sequence respectively.

Plots 3 & 6 are observed after receive amplification and fixed line equalization, using a PRBS-7 and PRBS-23 sequence respectively.

Note the decrease in vertical opening of the received eye using a PRBS-23 over PRBS-7. This indicates baseline wander in the receiver. This effect will be less pronounced with a 11th order polynomial scrambler, compared with the PRBS-23 sequence.

Plot 1:

PRBS 2⁷ - 1
50m "Horizontal" AT&T Systimax 1061B.
2 x 3m AMP patch cord.
Pair 2 activated.
Pair 4 activated.
Monitored after transmit transformer.



Plot 2:PRBS 2⁷ - 1

50m "Horizontal" AT&T Systimax 1061B.

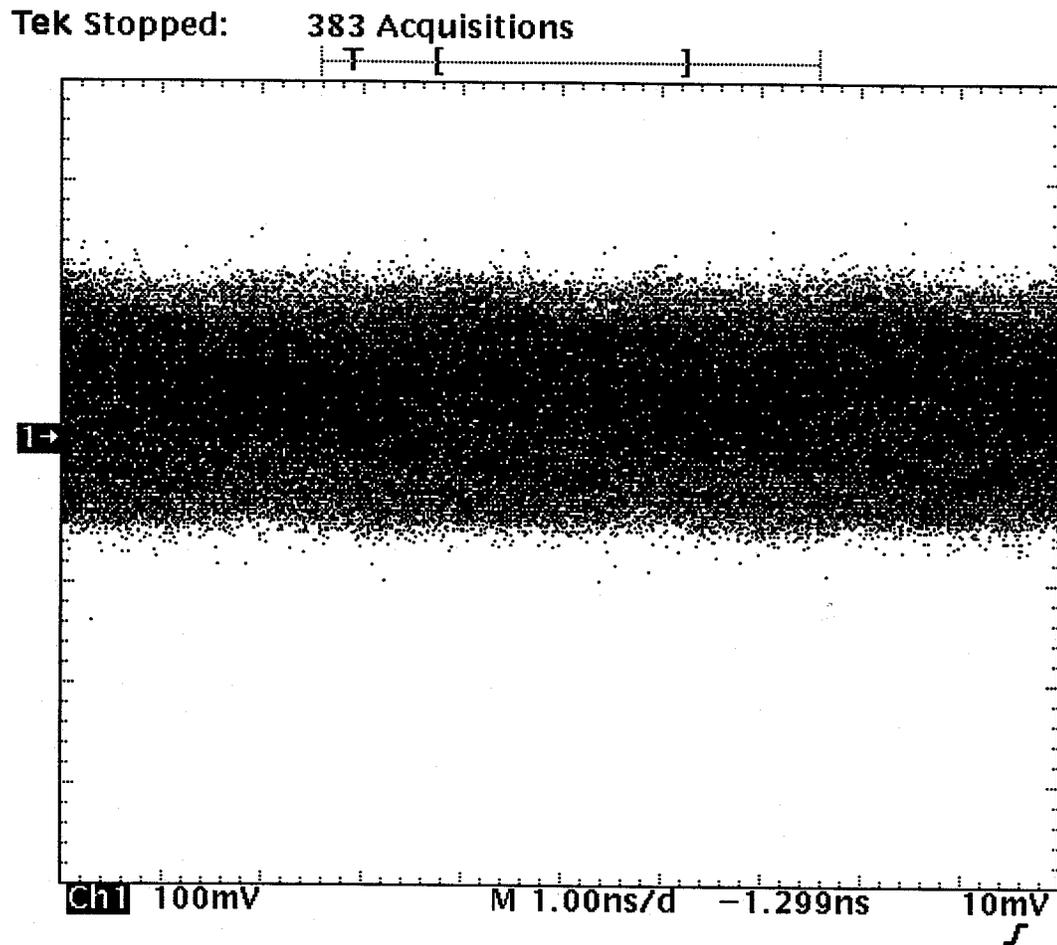
2 x 3m AMP patch cord.

Pair 2 activated.

Pair 4 activated.

Monitored after receive transformer (over 56m of cable).

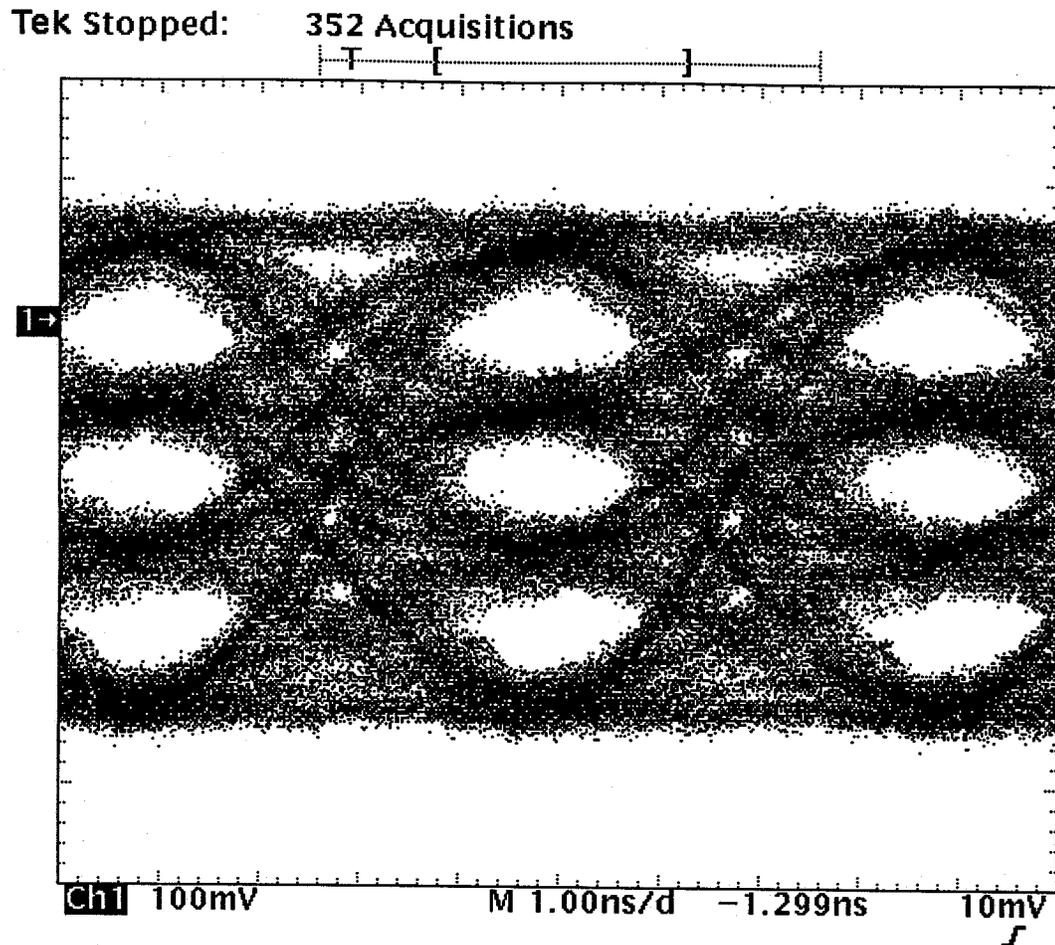
No amplification or equalization used.



Plot 3:

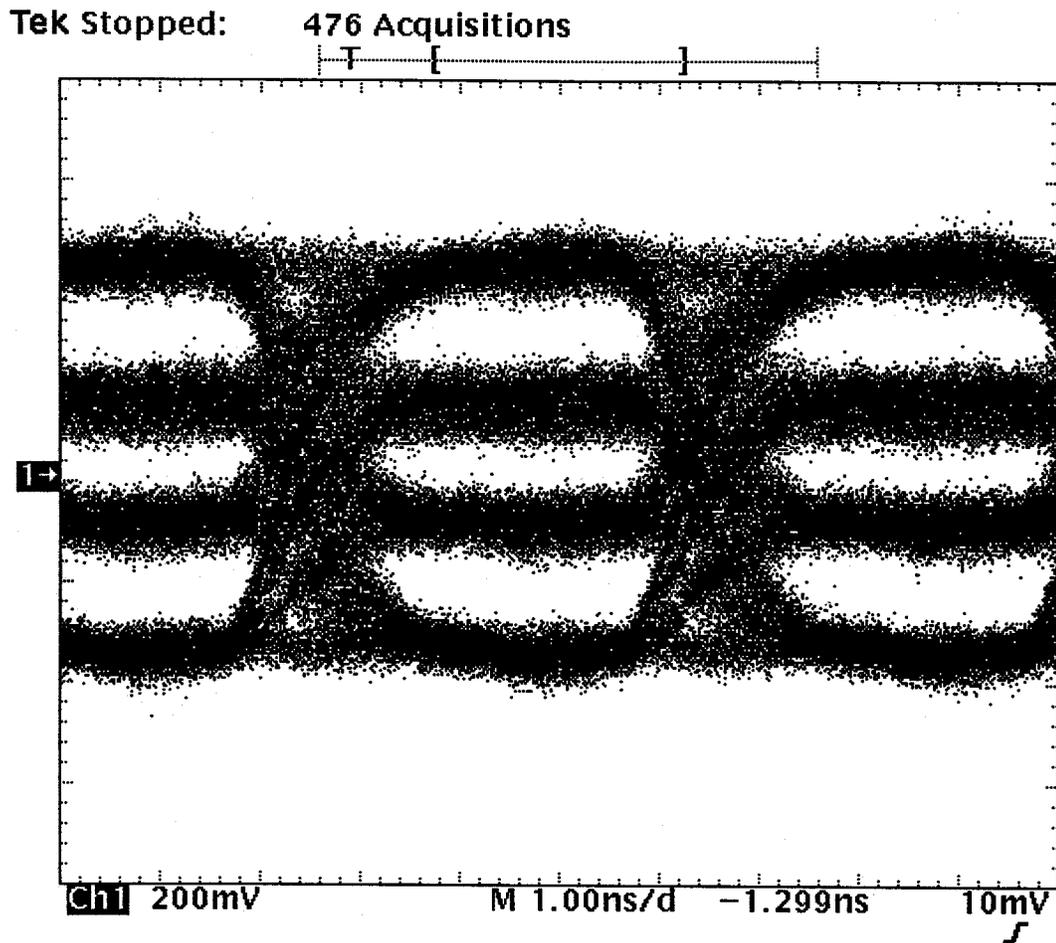
PRBS 2⁷ - 1
50m "Horizontal" AT&T Systimax 1061B.
2 x 3m AMP patch cord.
Pair 2 activated.
Pair 4 activated.
Monitored after receive pre-amplifier.
Simple fixed line equalization used.

This test ran error free for over 5 hours.
This corresponds to a BER better than 10⁻¹².



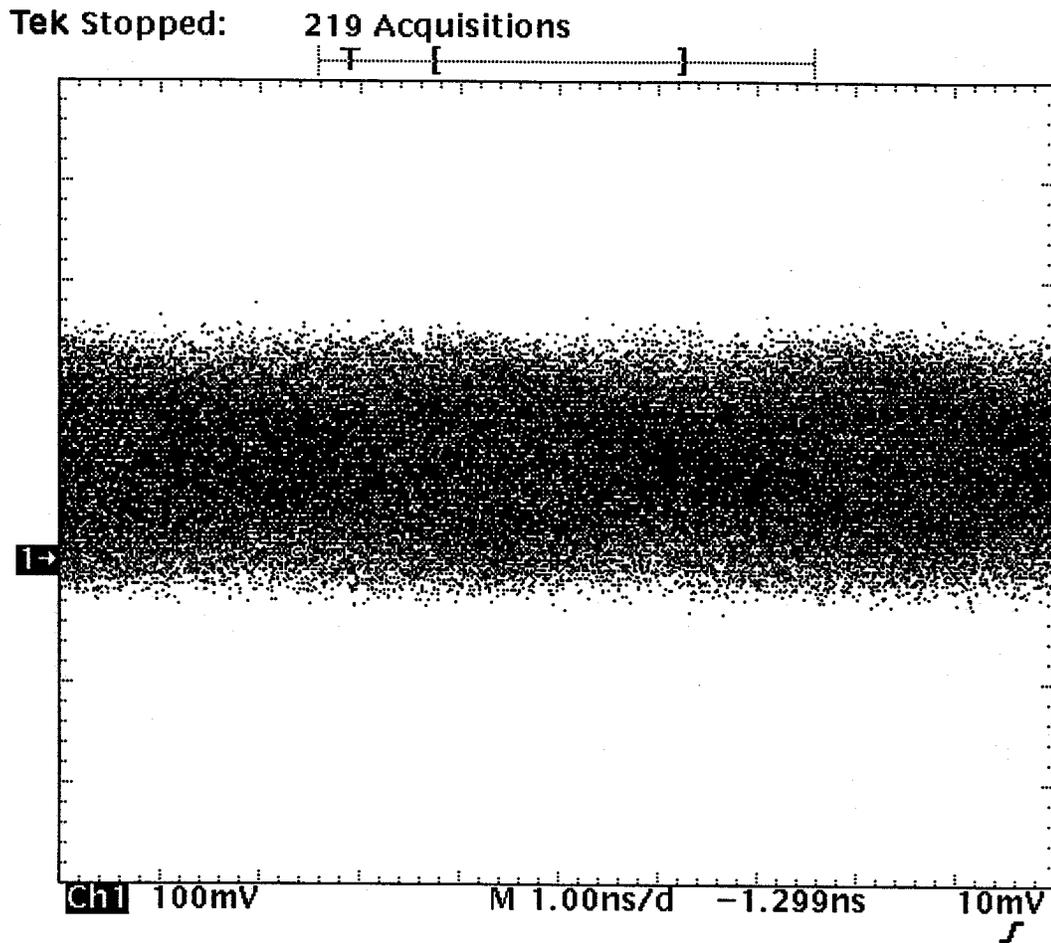
Plot 4:

PRBS 2²³ - 1
50m "Horizontal" AT&T Systimax 1061B.
2 x 3m AMP patch cord.
Pair 2 activated.
Pair 4 activated.
Monitored after transmit transformer.



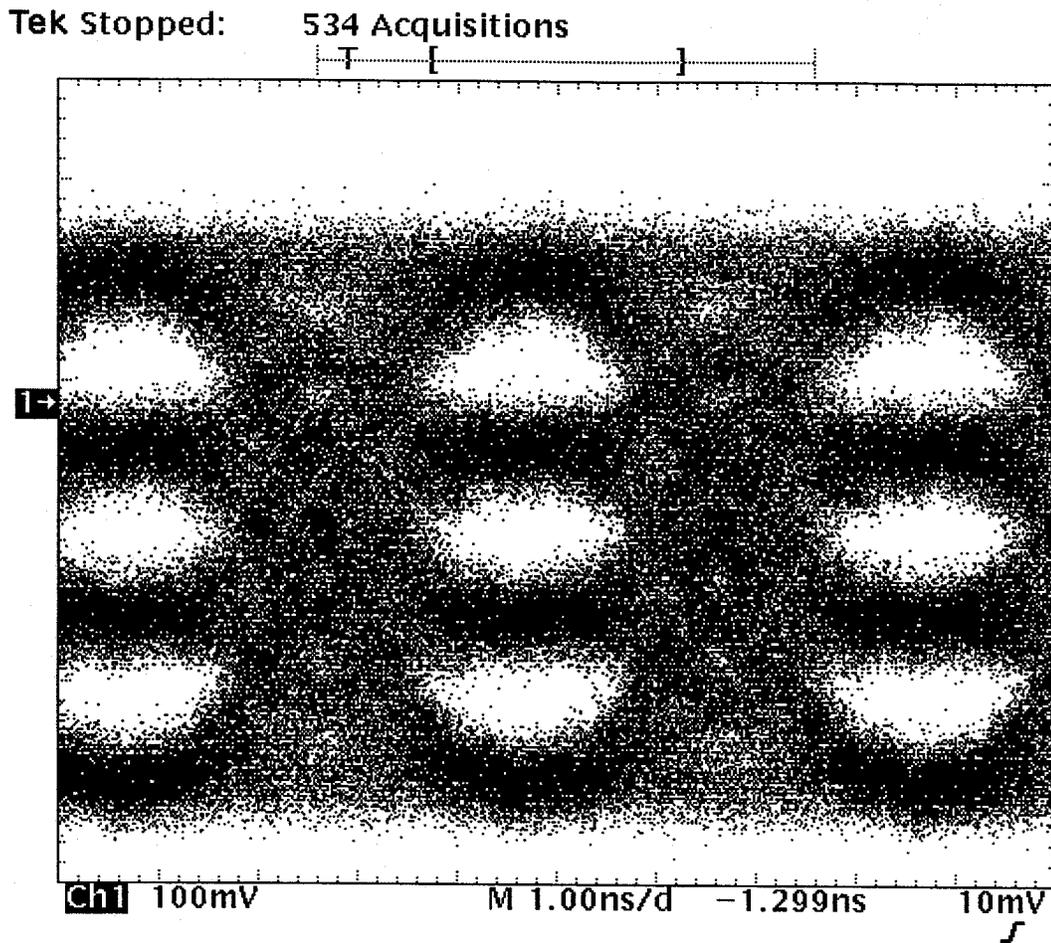
Plot 5:

PRBS 2²³ - 1
50m "Horizontal" AT&T Systimax 1061B.
2 x 3m AMP patch cord.
Pair 2 activated.
Pair 4 activated.
Monitored after receive transformer (over 56m of cable).
No amplification or equalization used.



Plot 6:

PRBS 2²³ - 1
50m "Horizontal" AT&T Systimax 1061B.
2 x 3m AMP patch cord.
Pair 2 activated.
Pair 4 activated.
Monitored after receive pre-amplifier.
Simple fixed line equalization used.



8. EMC Testing

8.1 Overview

In order to measure the EMC radiation of the cable when using this line coding operating at these frequencies over UTP-5 cable, an EMC test lab was used.

The remote outdoor facility provides excellent ambient conditions. The Equipment Under Test (EUT) included 2 evaluation boards, 50m of UTP-5 cable, a Bit Error Rate Tester (BERT) and appropriate power supplies. The EUT was placed on a table which was able to rotate 360 degrees. This allowed the worst case rotation to be taken for all frequency sweeps. Similarly, the antenna was able to rotate for both vertical and horizontal measurements, as well as being able to move vertically in order to measure the worst case radiation for different antenna heights. The test site was situated on a large metallic ground plane. All tests were controlled from a remote building, which provided full control of the antenna and EUT via buried control cables.

A diagram of the test facility is illustrated in Fig. 22.

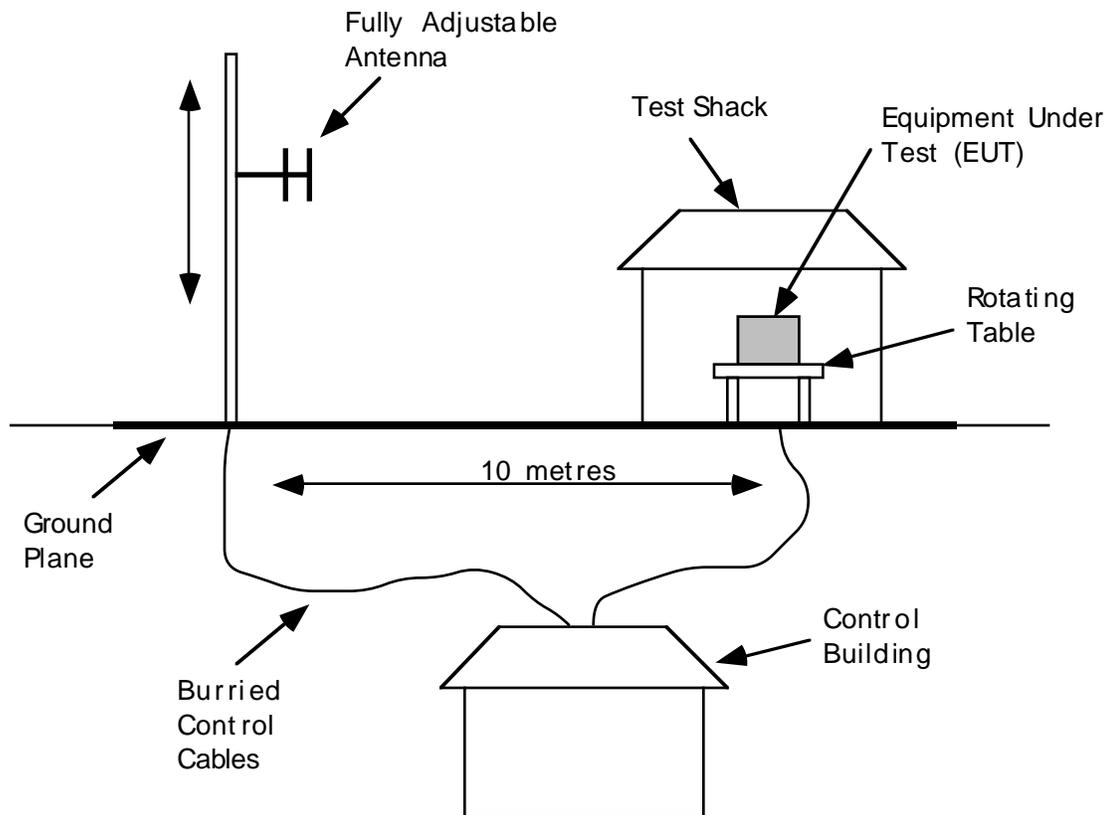


Fig. 22 EMC Test Facility

8.2 Test Setup

The equipment under test (EUT) used a configuration similar to configuration 3, described for BER performance in the previous section. This is illustrated in Fig. 23.

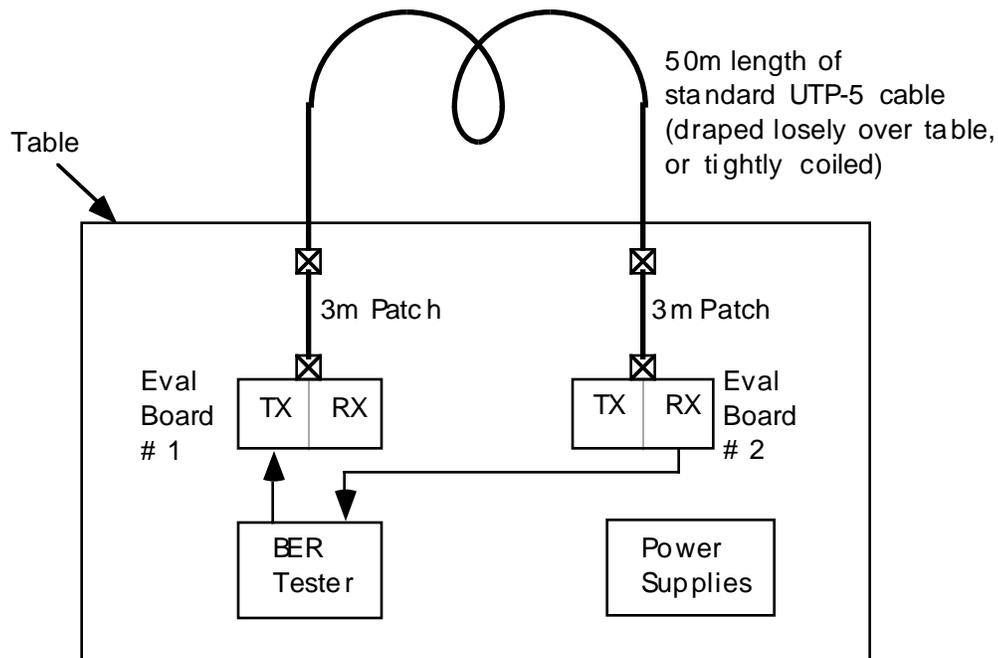


Fig. 23 EMC Test Configuration

As described earlier, it was possible to rotate the table holding the EUT in order to maximise the measured radiation for a given frequency spike.

Note that for all EMC tests with the exception of E1, shielding was provided around the BERT, both eval boards and the power supplies.

8.3 Test Procedure

Tests were run using a single 50m length of UTP-5 cable (AT&T Systimax 1061B).

The BER tester was used to provide 2 different data patterns (PRBS-7 and PRBS-23).

The individual tests are illustrated in Table 8:

Test Num	Test Overview	Line Code	Test Pattern	Cable Coiling	EMC Plot
E1	Ambient + BERT Tester	4-level	PRBS-7	tight	Plot E.1
E2	Ambient + 50m Cable	4-level	PRBS-23	loose	Plot E.2
E3	Ambient + 50m Cable	4-level	PRBS-7	loose	Plot E.3
E4	Ambient + 50m Cable	NRZ	PRBS-7	loose	Plot E.4

Table 8 Test Matrix for EMC Evaluation

8.4 Test Results

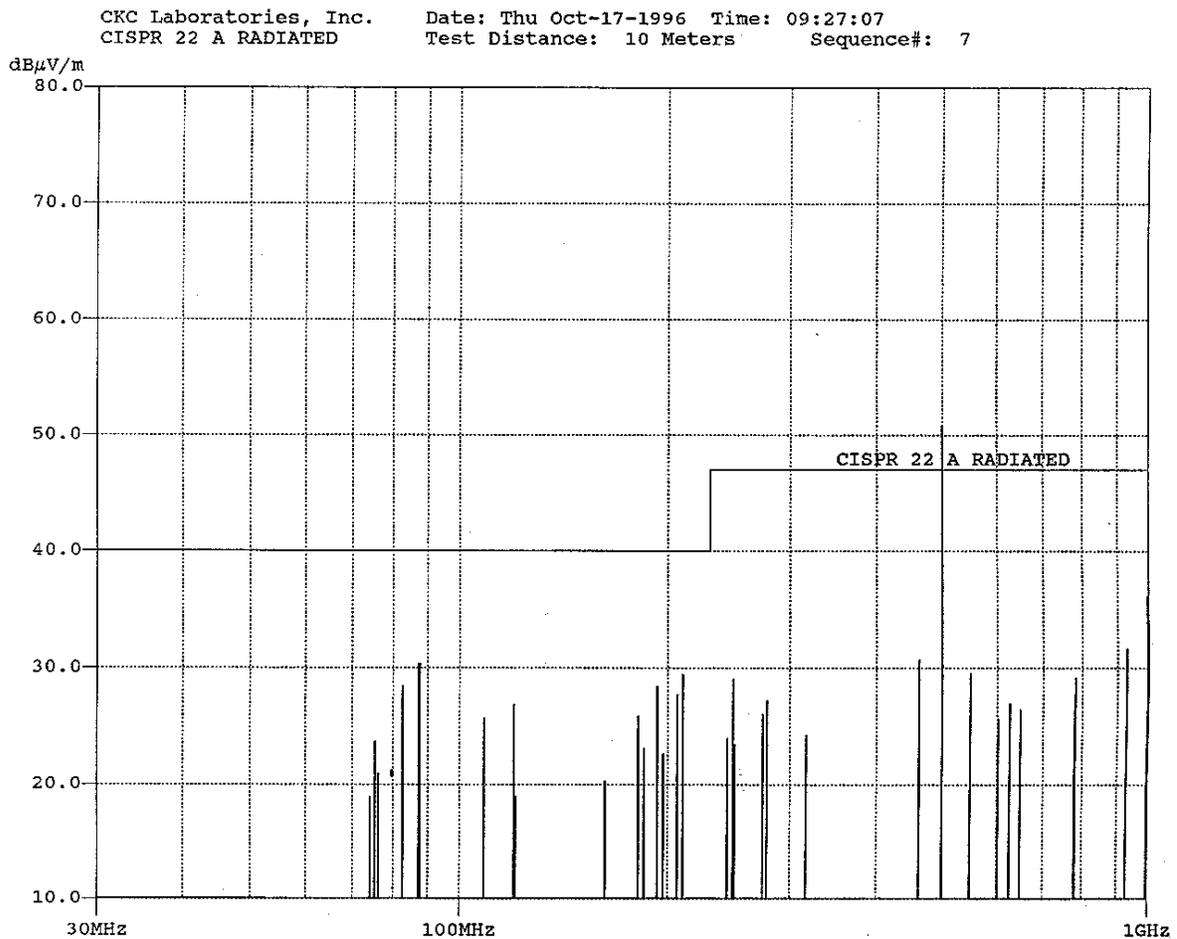
The measured EMC radiation for the tests above are illustrated in plots E1-E4.

Plot E.1 (Sequence #7)

All support on. EUT on. Short interconnect coax cables. Receiver and transmitter wrapped with foil. BERT generator PRBS-7 sending 500MHz NRZ signal to eval board #1.

Observations:

52dB spike at 500MHz due to BERT. All future tests to use BERT wrapped in foil. Illustrates the ambient EMC with the test equipment.



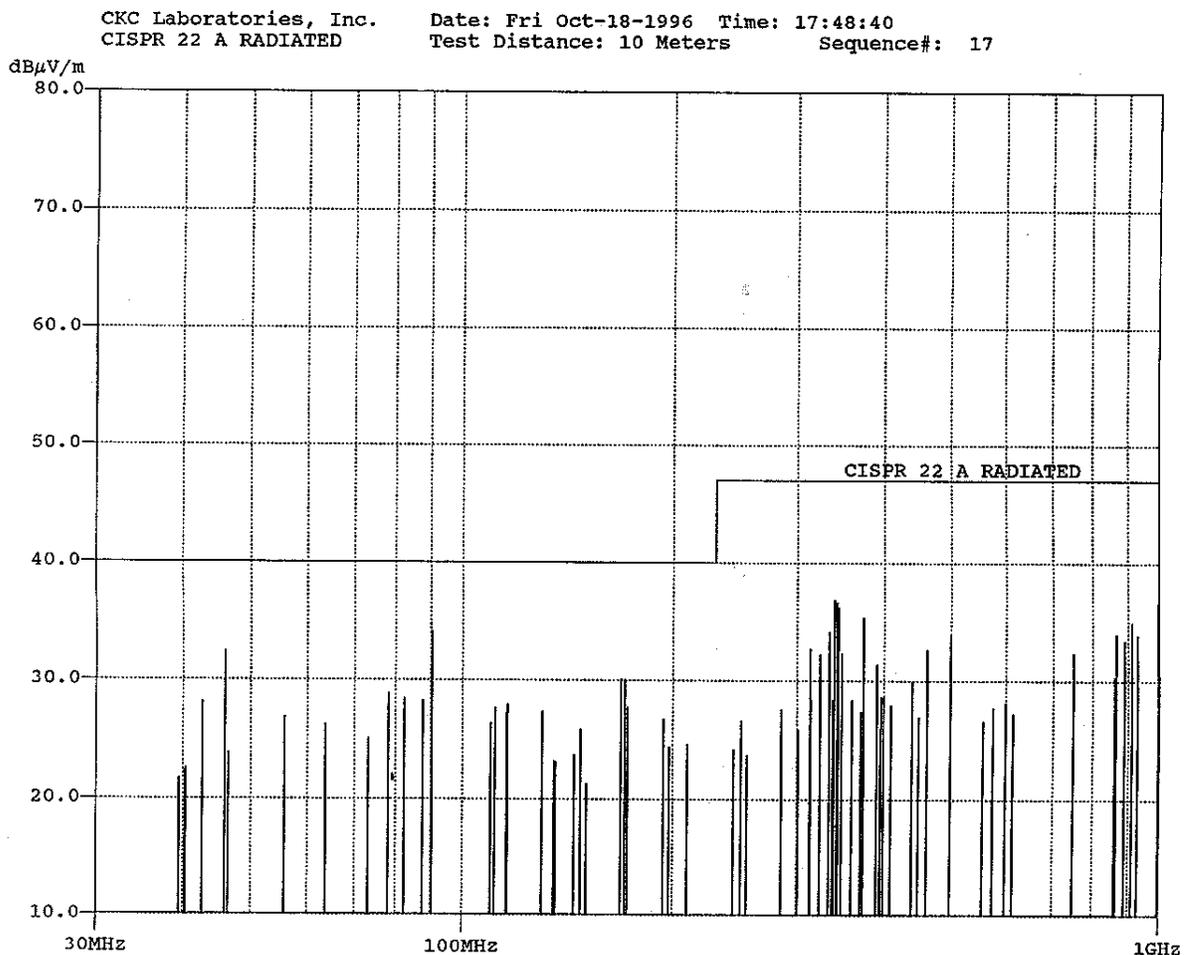
Plot. E.1 Sequence #7, Ambient + Unshielded BERT

Plot E.2 (Sequence #17)

All support on. EUT on. Short interconnect coax cables. Receiver and transmitter wrapped with foil. BERT generator PRBS-23 sending 500MHz NRZ signal to eval board #1. 56m UTP5 cable terminated to ground on PCB and wrapped around table ends. Transmit transformer ST6113 with choke. Receive transformer PE65508 no choke. Transformers internally shielded. Transmit output (0.85V with 30mV overshoot). Unused lines terminated and improved case/cable shielding.

Observations:

EMC passes with PRBS-23.



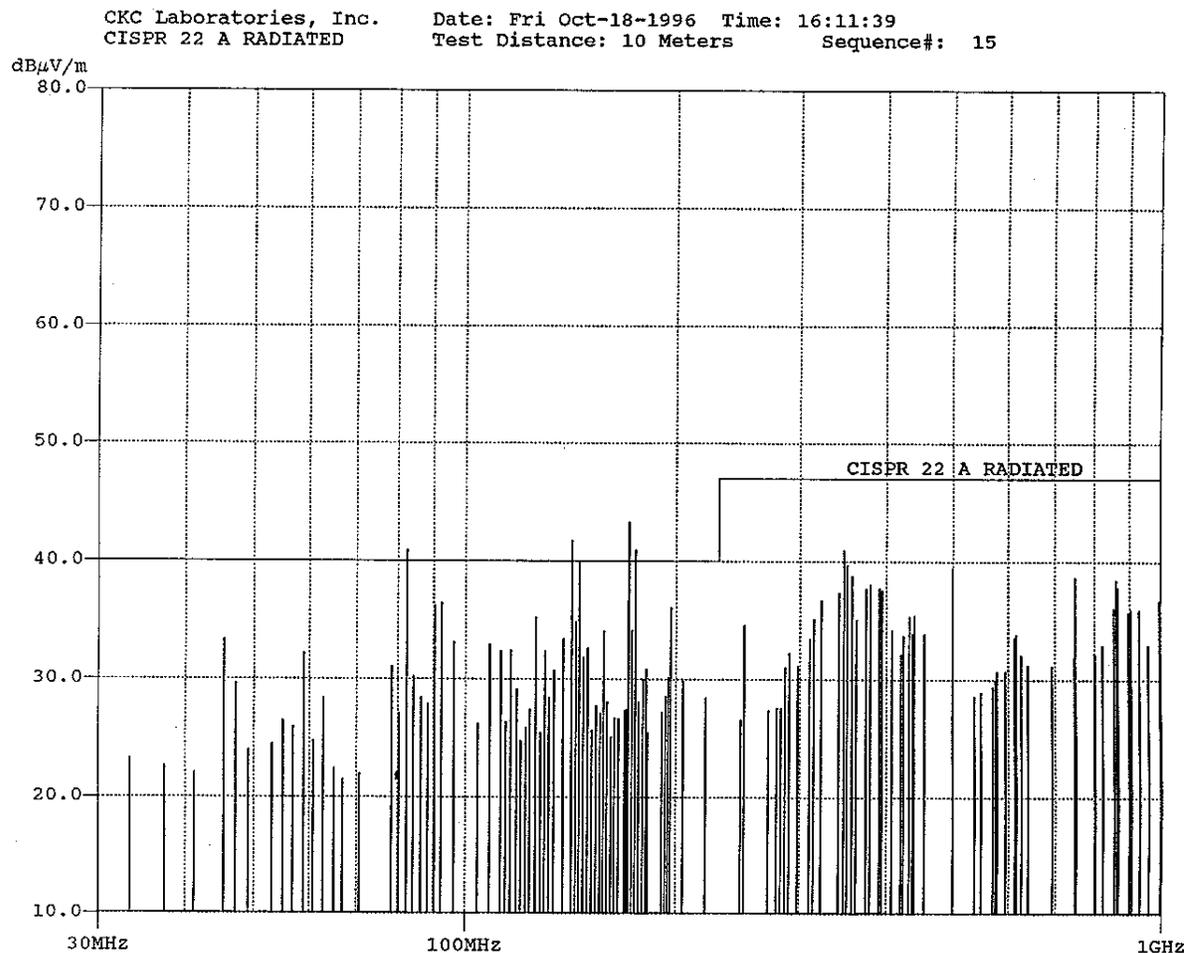
Plot. E.2 Sequence #17, 50m Cable, PRBS-23

Plot E.3 (Sequence #15)

All support on. EUT on. Short interconnect coax cables. Receiver and transmitter wrapped with foil. BERT generator PRBS-7 sending 500MHz NRZ signal to eval board #1. 56m UTP5 cable terminated to ground on PCB and wrapped around table ends. Transmit transformer ST6113 with choke. Receive transformer PE65508 no choke. Transformers internally shielded. Transmit output (0.9V with 50mV overshoot). Unused lines terminated and improved case shielding.

Observations:

EMC fails at various frequencies due to small run length (PRBS-7). Illustrates the requirement of scrambling to spread the spectral components.



Plot. E.3 Sequence #15, 50m Cable, PRBS-7

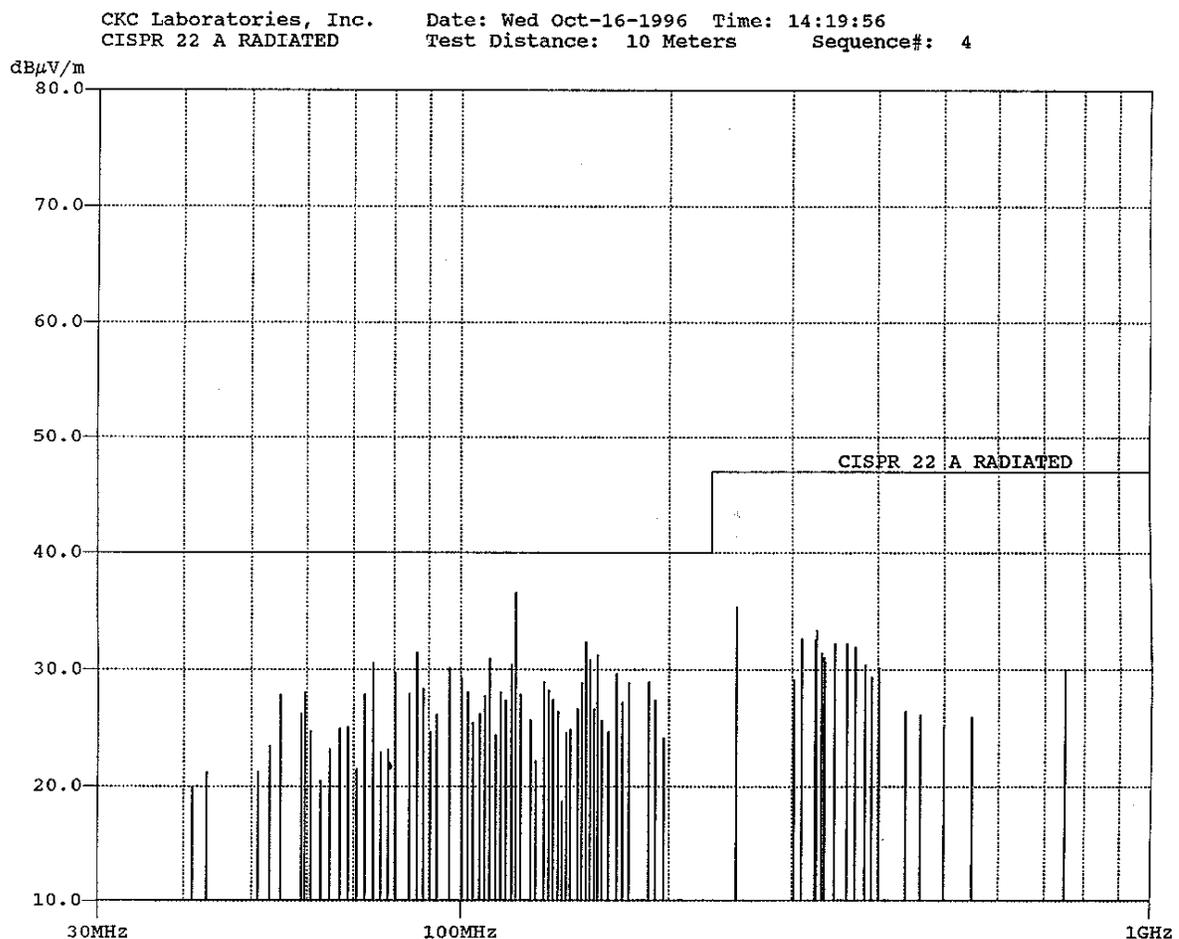
Plot E.4 (Sequence #4)

Similar setup as before but using NRZ directly from BERT, and not the 4-level coding, to drive the cable.

BERT driving 0.9V pk-pk 250MHz NRZ (PRBS-7) differential into transmit transformer (with choke), and then directly onto 56m of cable. Receive transformer (with choke) with 100Ω termination centre grounded.

Observations:

EMC passes with 250MHz NRZ.



Plot. E.4 Sequence #4, 50m Cable, 250MHz NRZ direct drive (no line coding)

9. Conclusion

In this report we have detailed a method for transmitting 1Gbps full duplex ethernet frames over a single UTP-5 cable bundle (4 pairs) using a 4LZS signaling scheme.

A set of escape sequences for frame delineation and control was presented that provide acceptable error robustness using an extended Hamming Code with a hamming distance of 4.

Empirical results were shown to demonstrate that acceptable performance could be achieved ($BER < 10^{-10}$) over distances of 50m. These empirical results were also correlated with simulation results.

Initial EMC results were provided which demonstrated that the method passes CISPR-22A radiated emissions tests, when using a PRBS-23 pattern as a data source. It was also shown that there is a requirement for frame scrambling, as indicated by the EMC failure when using PRBS-7 data patterns.

This method also lends itself for data transmission over longer distances or even higher rates using 2 cables containing a total of 8 pairs of UTP-5. Using 2 identical implementations operating in parallel, operation at 1Gbps using 2 cables would be achievable at distances of more than 100m.

Work for further study include the following:

- BER performance over temperature.
- Enhanced spectral and FCC measurements with frame scrambling.
- DC Balance requirements (decision feedback at receiver verses transmitter running digital sum correction).
- Transmit templates and slicing levels.

10. References

- [1] IEEE802.3u, "MAC Parameters, Physical Layer, Medium Attachment Units, and Repeaters for 100 Mbps Operation , Type 100BASE-T"
- [2] IEEE802.3x, "Specification for 802.3 Full-Duplex Operation"
- [3] EIA/TIA-568-A Standard, "Commercial Building & Wiring Telecommunications Wiring Standard", Letter Ballot, 1994.
- [4] J.J. Werner, "Tutorial on Carrierless AM/PM - part II", ANSI X3T9.5 TP-PMD submission, Austin, Feb1993.
- [5] T. Banwell, W. Stephens, "Line Code Selection for 155.52Mbps Data Transmission on Catagory 5 Cable Plant", IEEE JSAC, Vol 13, No. 9, Dec 1995.
- [6] S. Dabecki , "PCS Mapping Using 4LZS Codes", PMC-Sierra Inc, internal document - PMC-961007
- [7] R. Cam, "Construction of Simple and Provably Robust 4LZS Control Codes", PMC-Sierra Inc, internal document - PMC-961031
- [8] R. Cam, "Scrambling and 4LZS with Run Substitution", PMC-Sierra Inc, internal document - PMC-961021
- [9] ANSI X3.263-1995, "FDDI - Token Ring Twisted Pair Physical Layer Medium Dependent (TP-PMD)"