Information technology —
Telecommunications and information exchange between systems —
Local and metropolitan networks — specific requirements
Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications —

Amendment: Physical Layer and Management Parameters for 10 Gb/s Operation — Type 10GBASE-T

Sponsor

LAN MAN Standards Committee
of the
IEEE Computer Society

This Draft amendment to IEEE Std. 802.3 provides support to extend 10 Gb/s operation over ISO/IEC 11801:2002 Class E and Class F channels with a new Physical Layer (PHY) device. The PHY is known as 10GBASE-T. This draft D1.0 is being circulated as part of Task Force Ballot. The formal expiration date of this draft is October 1, 2004.

IEEE Standards Activities Department
Standards Licensing and Contracts
445 Hoes Lane, P.O. Box 1331
Piscataway, NJ 08855-1331, USA

Annex 28A

(normative)

Selector Field definitions

The Selector Field, S[4:0] in the Link Code Word, shall be used to identify the type of message being sent by Auto-Negotiation. The following table identifies the types of messages that may be sent. As new messages are developed, this table will be updated accordingly.

The Selector Field uses a 5-bit binary encoding, which allows 32 messages to be defined. All unspecified combinations are reserved. Reserved combinations shall not be transmitted.

Table 28A–1—Selector Field value mappings

S4	S3	S2	S1	S0	Selector description
0	0	0	0	0	Reserved for future Auto-Negotiation development
0	0	0	0	1	IEEE Std 802.3
0	0	0	1	0	IEEE Std 802.9 ISLAN-16T
0	0	0	1	1	IEEE Std 802.5
1	1	1	1	1	Reserved for future Auto-Negotiation development[a]

[a]For up-to-date information on the allocation of Auto-Negotiation Selector fields see http://www.ieee802.org/3/selectors/selectors.html

Annex 28B

(normative)

IEEE 802.3 Selector Base Page definition

This annex provides the Technology Ability Field bit assignments, Priority Resolution table, and Message Page transmission conventions relative to the IEEE 802.3 Selector Field value within the base page encoding.

As new IEEE 802.3 LAN technologies are developed, a reserved bit in the Technology Ability field may be assigned to each technology by the standards body.

The new technology will then be inserted into the Priority Resolution hierarchy and made a part of the Auto-Negotiation standard. The relative hierarchy of the existing technologies will not change, thus providing backward compatibility with existing Auto-Negotiation implementations.

It is important to note that the reserved bits are required to be transmitted as logic zeros. This guarantees that devices implemented using the current priority table will be forward compatible with future devices using an updated priority table.

28B.1 Selector field value

The value of the IEEE 802.3 Selector Field is S[4:0] = 00001.

28B.2 Technology Ability Field bit assignments

The Technology bit field consists of bits D5 through D12 (A0–A7, respectively) in the IEEE 802.3 Selector Base Page. Table 28B–1 summarizes the bit assignments.

Note that the order of the bits within the Technology Ability Field has no relationship to the relative priority of the technologies.

Setting Bit A5 or Bit A6 indicates that the DTE has implemented both the optional MAC control sublayer and the PAUSE function as specified in Clause 31 and Annex 31B. This capability is significant only when the link is configured for full duplex operation, regardless of data rate and medium. The encoding of Bits A5 and A6 is specified in Table 28B–2.

Table 28B–1—Technology Ability Field bit assignments

Bit	Technology	Minimum cabling requirement
A0	10BASE-T	Two-pair category 3
A1	10BASE-T full duplex	Two-pair category 3
A2	100BASE-TX	Two-pair category 5
A3	100BASE-TX full duplex	Two-pair category 5
A4	100BASE-T4	Four-pair category 3
A5	PAUSE operation for full duplex links	Not applicable
A6	Asymmetric PAUSE operation for full duplex Links	Not applicable
A7	Reserved for future technology	

Table 28B–2—Pause encoding

PAUSE (A5)	ASM_DIR (A6)	Capability
0	0	No PAUSE
0	1	Asymmetric PAUSE toward link partner
1	0	Symmetric PAUSE
1	1	Both Symmetric PAUSE and Asymmetric PAUSE toward local device

The PAUSE bit indicates that the device is capable of providing the symmetric PAUSE functions as defined in Annex 31B. The ASM_DIR bit indicates that asymmetric PAUSE is supported. The value of the PAUSE bit when the ASM_DIR bit is set indicates the direction the PAUSE frames are supported for flow across the link. Asymmetric PAUSE configuration results in independent enabling of the PAUSE receive and PAUSE transmit functions as defined by Annex 31B. See 28B.3 regarding PAUSE configuration resolution.

28B.3 Priority resolution

Since two devices may have multiple abilities in common, a prioritization scheme exists to ensure that the highest common denominator ability is chosen. The following list shall represent the relative priorities of the technologies supported by the IEEE 802.3 Selector Field value, where priorities are listed from highest to lowest.

a) 10GBASE-T full duplex
b) 1000BASE-T full duplex
c) 1000BASE-T
d) 100BASE-T2 full duplex
e) 100BASE-TX full duplex
f) 100BASE-T2
g) 100BASE-T4
h) 100BASE-TX
i) 10BASE-T full duplex
j) 10BASE-T

The rationale for this hierarchy is straightforward. 10BASE-T is the lowest common denominator and therefore has the lowest priority. Full duplex solutions are always higher in priority than their half duplex counterparts. 1000BASE-T has a higher priority than 100 Mb/s technologies. 100BASE-T2 is ahead of 100BASE-TX and 100BASE-T4 because 100BASE-T2 runs across a broader spectrum of copper cabling and can support a wider base of configurations. 100BASE-T4 is ahead of 100BASE-TX because 100BASE-T4 runs across a broader spectrum of copper cabling. The relative order of the technologies specified herein shall not be changed. As each new technology is added, it shall be inserted into its appropriate place in the list, shifting technologies of lesser priority lower in priority. If a vendor-specific technology is implemented, the priority of all IEEE 802.3 standard technologies shall be maintained, with the vendor specific technology inserted at any appropriate priority location.

The use of the PAUSE operation for full duplex links (as indicated by bits A5 and A6) is orthogonal to the negotiated data rate, medium, or link technology. The setting of these bits indicates the availability of additional DTE capability when full duplex operation is in use. The PAUSE function shall be enabled according to Table 28B–3 only if the Highest Common Denominator is a full duplex technology. There is no priority resolution associated with the PAUSE operation.

Table 28B–3—Pause resolution

Local device		Link partner		Local device resolution	Link partner resolution
PAUSE	**ASM_DIR**	**PAUSE**	**ASM_DIR**		
0	0	Don’t care	Don’t care	Disable PAUSE Transmit and Receive	Disable PAUSE Transmit and Receive
0	1	0	Don’t care	Disable PAUSE Transmit and Receive	Disable PAUSE Transmit and Receive
0	1	1	0	Disable PAUSE Transmit and Receive	Disable PAUSE Transmit and Receive
0	1	1	1	Enable PAUSE transmit Disable PAUSE receive	Enable PAUSE receive Disable PAUSE transmit
1	0	0	Don’t care	Disable PAUSE Transmit and Receive	Disable PAUSE Transmit and Receive
1	Don’t care	1	Don’t care	Enable PAUSE Transmit and Receive	Enable PAUSE Transmit and Receive
1	1	0	0	Disable PAUSE Transmit and Receive	Disable PAUSE Transmit and Receive
1	1	0	1	Enable PAUSE receive Disable PAUSE transmit	Enable PAUSE transmit Disable PAUSE receive

28B.4 Message Page transmission convention

Each series of Unformatted Pages shall be preceded by a Message Page containing a Message Code that defines how the following Unformatted Pages will be used.

Next Page message codes should be allocated globally across Selector Field values so that meaningful communication is possible between technologies using different Selector Field values.

Annex 28C

(normative)

Next Page Message Code Field definitions

The Message Code Field of a message page used in Next Page exchange shall be used to identify the meaning of a message. The following table identifies the types of messages that may be sent. As new messages are developed, this table will be updated accordingly.

The Message Code Field uses an 11-bit binary encoding that allows 2048 messages to be defined. All Message Codes not specified shall be reserved for IEEE use or allocation.

Table 28C–1—Message code field values

Message Code #	M 10	M 9	M 8	M 7	M 6	M 5	M 4	M 3	M 2	M 1	M 0	Message Code Description
0	0	0	0	0	0	0	0	0	0	0	0	Reserved for future Auto-Negotiation use
1	0	0	0	0	0	0	0	0	0	0	1	Null Message
2	0	0	0	0	0	0	0	0	0	1	0	One UP with Technology Ability Field follows
3	0	0	0	0	0	0	0	0	0	1	1	Two UPs with Technology Ability Field follow
4	0	0	0	0	0	0	0	0	1	0	0	One UP with Binary coded Remote fault follows
5	0	0	0	0	0	0	0	0	1	0	1	Organizationally Unique Identifier Tagged Message
6	0	0	0	0	0	0	0	0	1	1	0	PHY Identifier Tag Code
7	0	0	0	0	0	0	0	0	1	1	1	100BASE-T2 Technology Message Code. 100BASE-T2 Ability Page to follow using Unformatted Next Page
8	0	0	0	0	0	0	0	1	0	0	0	1000BASE-T Technology Message Code. Two 1000BASE-T Ability Pages to follow using Unformatted Next Pages.
9	0	0	0	0	0	0	0	1	0	0	1	10GBASE-T Technology Message Code. Two 10GBASE-T Ability Pages to follow using Unformatted Next Pages.
10.....	0	0	0	0	0	0	0	1	0	1	0	Reserved for future Auto-Negotiation use
.....2047	1	1	1	1	1	1	1	1	1	1	1	Reserved for future Auto-Negotiation use

28C.1 Message code #0—Auto-Negotiation reserved code 1

This code is reserved for future Auto-Negotiation function enhancements. Devices shall not transmit this code.

28C.2 Message code #1—Null Message code

The Null Message code shall be transmitted during Next Page exchange when the Local Device has no further messages to transmit and the Link Partner is still transmitting valid Next Pages. See 28.2.3.4 for more details.

28C.3 Message code #2—Technology Ability extension code 1

This Message Code is reserved for future expansion of the Technology Ability Field and indicates that a defined user code with a specific Technology Ability Field encoding follows.

28C.4 Message code #3—Technology Ability extension code 2

This Message Code is reserved for future expansion of the Technology Ability Field and indicates that two defined user codes with specific Technology Ability Field encodings follow.

28C.5 Message code #4—Remote fault number code

This Message Code shall be followed by a single user code whose encoding specifies the type of fault that has occurred. The following user codes are defined:

0: RF Test
This code can be used to test Remote Fault operation.
1: Link Loss
2: Jabber
3: Parallel Detection Fault
This code may be sent to identify when bit 6.4 is set.

28C.6 Message code #5—Organizationally Unique Identifier (OUI) tag code

The OUI Tagged Message shall consist of a single message code of 0000 0000 0101 followed by four user codes defined as follows. The first user code shall contain the most significant 11 bits of the OUI (bits 23:13) with the most significant bit in bit 10 of the user code. The second user code shall contain the next most significant 11 bits of the OUI (bits 12:2) with the most significant bit in bit 10 of the user code. The third user code shall contain the remaining least significant 2 bits of the OUI (bits 1:0) with the most significant bit in bit 10 of the user code. Bits 8:0 of the fourth user contain a user-defined user code value that is specific to the OUI transmitted. The fourth and final user code shall contain a user-defined user code value that is specific to the OUI transmitted.

28C.7 Message code #6—PHY identifier tag code

The PHY ID tag code message shall consist of a single message code of 0000 0000 0110 followed by four user codes defined as follows. The first user code shall contain the most significant 11 bits of the PHY ID

(2.15:5) with the most significant bit in bit 10 of the user code. The second user code shall contain bits 2.4:0 to 3.15:10 of the PHY ID with the most significant bit in bit 10 of the user code. The third user code shall contain bits 3.9:0 of the PHY ID with the most significant bit in bit 10 of the user code. Bit 0 in the third user code shall contain a user-defined user code value that is specific to the PHY ID transmitted. The fourth and final user code shall contain a user-defined user code value that is specific to the PHY ID transmitted.

28C.8 Message code #2047—Auto-Negotiation reserved code 2

This code is reserved for future Auto-Negotiation function enhancements. Devices shall not transmit this code.

28C.9 Message code #7—100BASE-T2 technology message code

Clause 32 (100BASE-T2) uses Next Page Message Code 7 to indicate that T2 implementations will follow the transmission of this page [the initial, Message (formatted) Next Page] with two unformatted Next Pages which contain information defined in 32.5.4.2.

28C.10 Message Code #8 - 1000BASE-T technology message code

Clause 40 (1000BASE-T) uses Next Page Message Code 8 to indicate that 1000BASE-T implementations will follow the transmission of this page [the initial, Message (formatted) Next Page] with two unformatted Next Pages that contain information defined in 40.5.1.2.

28C.11 Message Code #9 - 10GBASE-T technology message code

#CrossRef# Clause 55 (10GBASE-T) uses Next Page Message Code 9 to indicate that 10GBASE-T implementations will follow the transmission of this page [the initial, Message (formatted) Next Page] with two unformatted Next Pages that contain information defined in #CrossRef# TBD.

Annex 28D

(normative)

Description of extensions to Clause 28 and associated annexes

28D.1 Introduction

This annex is to be used to document extensions and modifications to Clause 28 required by IEEE 802.3 clauses and other standards that use Auto-Negotiation and that were approved after June 1995. It provides a single location to define such extensions and modifications without changing the basic contents of Clause 28.

Subclause 28D.2 lists those clauses and standards that require extensions to Clause 28 and provides pointers to the subclauses where those extensions are listed.

28D.2 Extensions to Clause 28

28D.2.1 Extensions required for Clause 31 (full duplex)

Clause 31 (full duplex) requires the use of Auto-Negotiation. Extensions to Clause 28 and associated annexes required for the correct operation of full duplex are shown in 28D.3.

28D.2.2 Extensions required for Clause 32 (100BASE-T2)

Clause 32 (100BASE-T2) requires the use of Auto-Negotiation. Extensions to Clause 28 required for correct operation of 100BASE-T2 are shown in 28D.4.

28D.3 Extensions for Clause 31

Full duplex requires the use of bit A5 in the Technology Ability Field of the IEEE 802.3 Selector Base Page. (This bit is also defined as MII bit 4.10.) This bit was previously defined as “reserved for future technology.”

Bit	Technology	Minimum cabling requirement
A5	PAUSE operation for full duplex links	Not applicable

Bit A5 (PAUSE operation for full duplex links) signifies that the DTE has implemented both the optional MAC Control sublayer and the PAUSE function as specified in Clause 31 and Annex 31B. This capability is significant only when the link is configured for full duplex operation, regardless of data rate and medium.

28D.4 Extensions for Clause 32 (100BASE-T2)

Clause 32 (100BASE-T2) makes special use of Auto-Negotiation and requires additional MII registers. This use is summarized below. Details are provided in 32.5.

Auto-Negotiation is mandatory for 100BASE-T2 (32.1.3.4).

100BASE-T2 introduces the concept of MASTER and SLAVE to define DTEs and to facilitate the timing of transmit and receive operations. Auto-Negotiation is used to provide information used to configure MASTER/SLAVE status (32.5.4.3).

100BASE-T2 uses unique next page transmit and receive registers (MII Registers 8, 9 and 10) in conjunctions with Auto-Negotiation. These registers are in addition to Registers 0–7 as defined in 28.2.4 (32.5.2).

100BASE-T2 use of Auto-Negotiation generates information which is stored in configuration and status bits defined for the MASTER-SLAVE Control register (MII Register 9) and the MASTER-SLAVE Status register (MII Register 10).

100BASE-T2 requires an ordered exchange of next page messages (32.5.1).

100BASE-T2 parameters are configured based on information provided by the ordered exchange of next page messages.

100BASE-T2 adds new message codes to be transmitted during Auto-Negotiation (32.5.4.2).

100BASE-T2 adds 100BASE-T2 full duplex and half duplex capabilities to the priority resolution table (28B.3) and MII Status Register (22.2.4.2).

T2 is defined as a valid value for “x” in 28.3.1 (e.g., link_status_T2). T2 represents that the 100BASE-T2 PMA is the signal source.

28D.5 Extensions required for Clause 40 (1000BASE-T)

Clause 40 (1000BASE-T) makes special use of Auto-Negotiation and requires additional MII registers. This use is summarized below. Details are provided in 40.5.

a) Auto-Negotiation is mandatory for 1000BASE-T. (40.5.1)
b) 1000BASE-T requires an ordered exchange of Next Page messages. (40.5.1.2)
c) 1000BASE-T parameters are configured based on information provided by the ordered exchange of NEXT Page messages.
d) 1000BASE-T uses MASTER and SLAVE to define PHY operations and to facilitate the timing of transmit and receive operations. Auto-Negotiation is used to provide information used to configure MASTER-SLAVE status.(40.5.2)
e) 1000BASE-T transmits and receives Next Pages for exchange of information related to MASTER-SLAVE operation. The information is specified in MII registers 9 and 10 (see 32.5.2 and 40.5.1.1), which are required in addition to registers 0-8 as defined in 28.2.4.
f) 1000BASE-T adds new message codes to be transmitted during Auto-Negotiation. (40.5.1.3)
g) 1000BASE-T adds 1000BASE-T full duplex and half duplex capabilities to the priority resolution table. (28B.3) and MII Extended Status Register (22.2.2.4)
h) 1000BASE-T is defined as a valid value for “x” in 28.3.1 (e.g., link_status_1GigT.) 1GigT represents that the 1000BASE-T PMA is the signal source.
i) 1000BASE-T supports Asymmetric Pause as defined in Annex 28B.

28D.6 Extensions required for #CrossRef# Clause 55 (10GBASE-T)

#CrossRef# Clause 55 (10GBASE-T) makes special use of Auto-Negotiation and requires additional MII registers. This use is summarized below. Details are provided in #CrossRef# TBD.

a) Auto-Negotiation is mandatory for 10GBASE-T. (#CrossRef# TBD)
b) 10GBASE-T requires an ordered exchange of Next Page messages. (#CrossRef# TBD)
c) 10GBASE-T parameters are configured based on information provided by the ordered exchange of NEXT Page messages.
d) 10GBASE-T uses MASTER and SLAVE to define PHY operations and to facilitate the timing of transmit and receive operations. Auto-Negotiation is used to provide information used to configure MASTER-SLAVE status.(#CrossRef# TBD)
e) 10GBASE-T transmits and receives Next Pages for exchange of information related to MASTER-SLAVE operation. The information is specified in MII registers 9 and 10 (see 32.5.2, 40.5.1.1, and #CrossRef# TBD), which are required in addition to registers 0-8 as defined in 28.2.4.
f) 10GBASE-T adds new message codes to be transmitted during Auto-Negotiation. (#CrossRef# TBD)
g) 10GBASE-T adds 10GBASE-T full duplex capabilities to the priority resolution table. (28B.3) and TBD (#CrossRef# TBD)
h) 10GBASE-T is defined as a valid value for “x” in 28.3.1 (e.g., link_status_10GigT.) 10GigT represents that the 10GBASE-T PMA is the signal source.
i) 10GBASE-T supports Asymmetric Pause as defined in Annex 28B.
j) TBD

55. Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) sublayer and baseband medium, type 10GBASE-T

55.1 Overview

The 10GBASE-T PHY is one of the Gigabit Ethernet family of high-speed CSMA/CD network specifications. The 10GBASE-T Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) and baseband medium specifications are intended for users who want 10 Gb/s performance over balanced twisted-pair structured cabling systems. 10GBASE-T signaling requires four pairs of balanced cabling, as specified in ISO/IEC 11801 Edition 2 with appropriate augmentation as specified in Clause 55.7.

This clause defines the type 10GBASE-T PCS, type 10GBASE-T PMA sublayer, and type 10GBASE-T Medium Dependent Interface (MDI). Together, the PCS and the PMA sublayer comprise a 10GBASE-T Physical layer (PHY). Fully functional, electrical, and mechanical specifications for the type 10GBASE-T PCS, PMA, and MDI are provided in this document. This clause also specifies the baseband medium used with 10GBASE-T.

55.1.1 Objectives

The following are the objectives of 10GBASE-T:

a) Support full duplex operation at 10 Gb/s over four connector, twisted-pair copper cabling with distances of:
   1) At least 100 m on four pair Class F balanced copper cabling as specified in ISO/IEC 11801 Edition 2.
   2) At least 55 m on four pair Class E balanced copper cabling as specified in ISO/IEC 11801 Edition 2 (with frequency extrapolation of the Class E specifications as defined in Clause 55.7).
   3) 100 m on four pair Class E-augmented balanced copper cabling as specified in ISO/IEC 11801 Edition 2 (with augmentation as defined in Clause 55.7).

b) Preserve the 802.3/Ethernet frame format at the MAC Client service Interface

c) Preserve minimum and maximum frame size of the current 802.3 Standard

d) Support Auto-Negotiation (Clause 28)

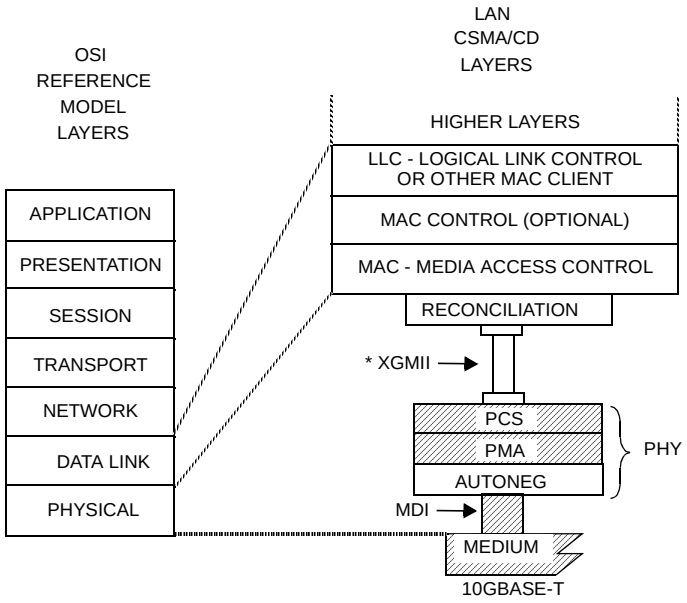
e) Meet CISPR/FCC Class A operation

f) Support a Bit Error Rate of less than or equal to 10^{-12} on all supported distances and Classes

Editor's Note: When the extrapolation of the Class E specifications specified in Clause 55.7 is incorporated into the appropriate TIA/ISO/IEC specifications, we will pull in references to these in here. Frequency extrapolation of the Class E specification is already available in the TR-42 Draft Technical Report entitled "Assessment of installed class E and class F cabling performance beyond their maximum specified frequencies.

55.1.2 Relationship of 10GBASE-T to other standards

Relations between the 10GBASE-T PHY, the ISO Open Systems Interconnection (OSI) Reference Model, and the IEEE 802.3 CSMA/CD LAN Model are shown in Figure 55–1. The PHY sub-layers (shown shaded) in Figure 55–1 connect one Clause 4 Media Access Control (MAC) layer to the medium.



MDI = MEDIUM DEPENDENT INTERFACE
XGMII = TEN GIGABIT MEDIA INDEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PHY = PHYSICAL LAYER DEVICE

*XGMII is optional.

Figure 55–1—Type 10GBASE-T PHY relationship to the ISO Open Systems Interconnection (OSI) Reference Model and the IEEE 802.3 CSMA/CD LAN Model

55.1.3 Operation of 10GBASE-T

The 10GBASE-T PHY employs full duplex baseband transmission over four pairs of balanced cabling. The aggregate data rate of 10 Gb/s is achieved by transmission at a data rate of 2500 Mb/s over each wire pair, as shown in Figure 55–2. The use of hybrids and cancellers enables full duplex transmission by allowing symbols to be transmitted and received on the same wire pairs at the same time. Baseband signaling with a modulation rate of TBD Msymbols/s is used on each of the wire pairs. The transmitted symbols are selected from a four-dimensional TBD-level symbol constellation. Each four-dimensional symbol can be viewed as a 4-tuple (A_n, B_n, C_n, D_n) of one-dimensional symbols taken from the set {TBD}. 10GBASE-T uses a continuous signaling system; in the absence of data, control symbols are transmitted. Data and Control symbols are embedded in a framing scheme which runs continuously after startup of the link. TBD-level Pulse Amplitude Modulation (PAMTBD) is employed for transmission over each wire pair. The modulation symbol rate of TBD Msymbols/s results in a symbol period of TBD ns.

A 10GBASE-T PHY can be configured either as a MASTER PHY or as a SLAVE PHY. The MASTER-SLAVE relationship between two stations sharing a link segment is established during Auto-Negotiation (see TBD). The MASTER PHY uses a local clock to determine the timing of transmitter operations. The SLAVE PHY recovers the clock from the received signal and uses it to determine the timing of transmitter operations, i.e., it performs loop timing, as illustrated in Figure 55–3. In a multiport to single-

port connection, the multiport device is typically set to be MASTER and the single-port device is set to be SLAVE.

The PCS and PMA subclauses of this document are summarized in 55.1.3.1 and 55.1.3.2. Figure 55–3 shows the functional block diagram.

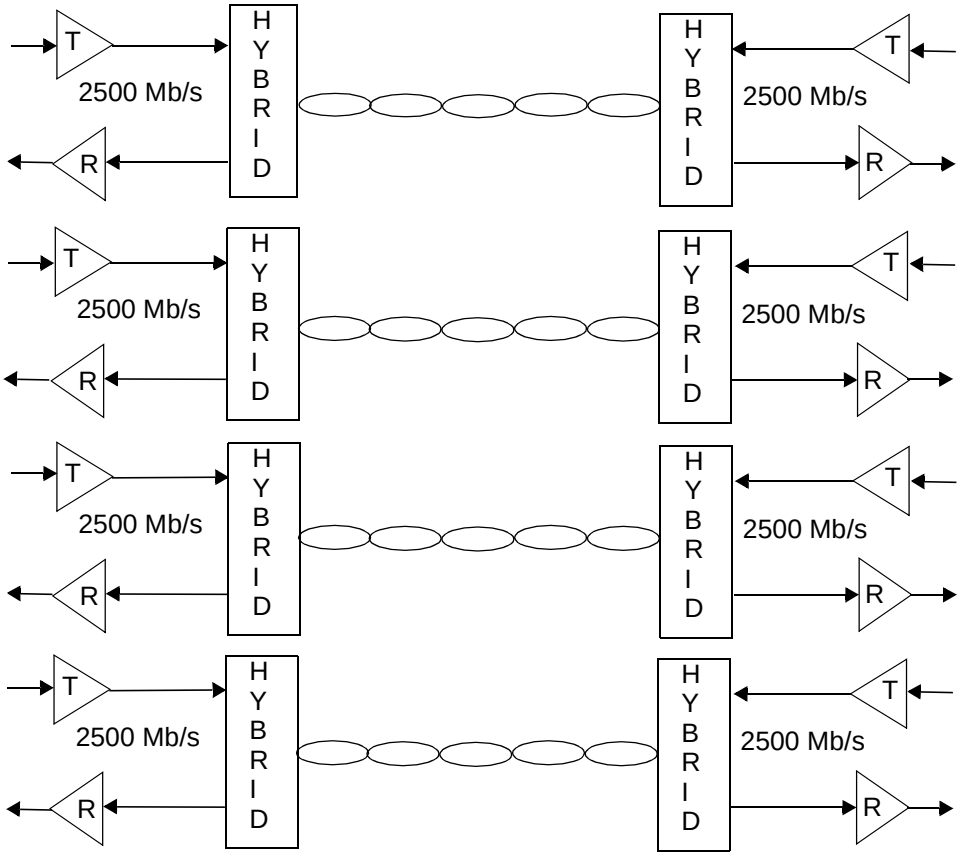


Figure 55–2—10GBASE-T topology

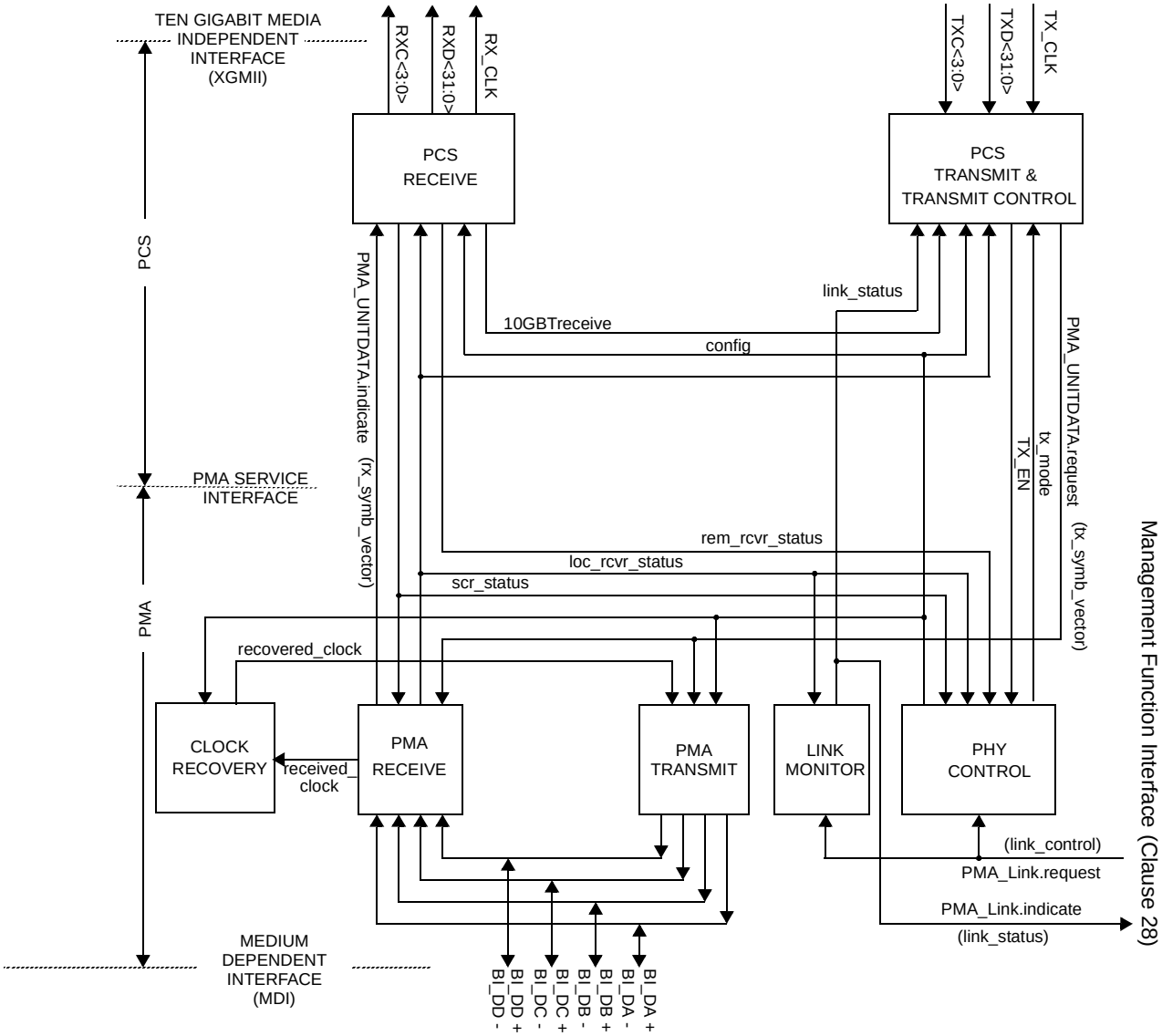


Figure 55–3—Functional block diagram

NOTE—The recovered_clock arc is shown to indicate delivery of the received clock signal back to PMA TRANSMIT for loop timing.

55.1.3.1 Physical Coding Sublayer (PCS)

The 10GBASE-T PCS couples a Ten Gigabit Media Independent Interface (XGMII), as described in Clause 46, to the 10GBASE-T Physical Medium Attachment (PMA) sublayer.

In the transmit direction, in normal mode, the PCS takes eight XGMII data octets provided by two consecutive transfers of XGMII service interface on TXD<31:0> and groups them into 64-bit blocks with the 64-bit block boundaries aligned with the XGMII packet boundaries as indicated by the XGMII transmit control signals (TXCn = 1). The PCS then adds 1 control bit to the head of this block and then scrambles all 65 bits. The resulting 65-bit blocks are then passed to the LDPC encoder to generate encoded blocks of TBD four dimensional signals. An LDPC frame sync header of TBD four dimensional signals is then added to the head of TBD LDPC block(s) to generate a continuous stream of four dimensional symbols that is passed on the PMA via PMA_UNITDATA.request signal. The PMA transmit block operates continuously this stream of four dimensional symbols. Details of the mapping are covered in TBD.

In the receive direction, in normal mode, the PCS processes code-groups received from the remote PHY via the PMA in TBD bit blocks and maps them to the XGMII service interface in the receive path. In this receive processing scheme, symbol clock synchronization is done by the PMA receive function.

In addition to the normal mode of operation, the PCS supports TBD additional modes. Furthermore, the PCS contains a management interface.

The PCS functions and state diagrams are specified in 55.3. The signals provided by the PCS at the XGMII conform to the interface requirements of Clause 46. The PCS Service Interfaces to the XGMII and the PMA are abstract message-passing interfaces specified in 55.2.

55.1.3.2 Physical Medium Attachment (PMA) sublayer

The PMA couples messages from the PCS service interface onto the balanced cabling physical medium via the Medium Dependent Interface (MDI) and provides the link management and PHY Control functions. The PMA provides full duplex communications at TBD Msymbols/s over four pairs of balanced cabling up to 100 m in length.

The PMA Transmit function comprises four transmitters to generate continuous time analog signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD, as described in 55.4.3.1. In normal mode, each four dimensional symbol received from the PCS transmit function undergoes multiple stages of processing. First the symbol goes through a Tomlinson Harashima Precoder (THP) which maps the M-ary PAM input in each dimension of the four dimensional symbol into a quasi-continuos discrete time value in the range {TBD, -TBD}. This THP processed four dimensional symbol stream is then processed through a transmit shaping filter and is then passed on to four digital to analog converters (DAC). The DAC outputs may be further processed with continuos time filters to roll off the high frequency spectral response to limit high frequency emissions and are then applied to the four balanced pairs via the MDI port.

The PMA Receive function comprises four independent receivers for pulse-amplitude modulated signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD, as described in 55.4.3.2. The receivers are responsible for acquiring symbol timing and, when operating in normal mode, for cancelling echo, near end cross talk, far end cross talk and equalizing the signal. The 4-D symbols are provided to the PCS receive function via the PMA_UNITDATA.indicate message. The PMA also contains functions for Link Monitor.

The PMA PHY Control function generates signals that control the PCS and PMA sublayer operations. PHY Control begins following the completion of Auto-Negotiation and provides the start-up functions required for successful 10GBASE-T operation. It determines whether the PHY operates in a normal state, enabling data transmission over the link segment, or whether the PHY sends special code-groups that are used in the

TBD mode(s). The latter occurs when either one or both of the PHYs that share a link segment are not operating reliably.

PMA functions and state diagrams are specified in 55.4. PMA electrical specifications are given in 55.5.

55.1.4 Signaling

10GBASE-T signaling is performed by the PCS generating continuous code-group sequences that the PMA transmits over each wire pair. The signaling scheme achieves a number of objectives including

a) Forward Error Correction (FEC) coded symbol mapping for data.
b) Algorithmic mapping from TXD<31:0> and TXC<3:0> to four dimensional symbols in the transmit path.
c) Algorithmic mapping from the received four dimensional signals on the MDI port to RXD<31:0> and RXC<3:0> on the XGMII interface.
d) Uncorrelated symbols in the transmitted symbol stream.
e) No correlation between symbol streams traveling both directions on any pair combination.
f) No correlation between symbol streams on pairs BI_DA, BI_DB, BI_DC, and BI_DD.
g) Block framing and other control signals.
h) Ability to signal the status of the local receiver to the remote PHY to indicate that the local receiver is not operating reliably and requires retraining.
i) Ability to automatically detect and correct for pair swapping and unexpected crossover connections.
j) Ability to automatically detect and correct for incorrect polarity in the connections.
k) Ability to automatically correct for differential delay variations across the wire-pairs.

The PHY operates in two basic modes, normal mode or training mode. In normal mode, PCS generates a continuous stream of four dimensional symbols that are transmitted via the PMA at TBD power levels. In training mode, the PCS is directed to generate only TBD symbols for transmission by the PMA, which enable the receiver at the other end to train until it is ready to operate in normal mode. (See the PCS reference diagram in 55.2.)

55.1.5 Inter-sublayer interfaces

All implementations of the balanced cabling link are compatible at the MDI. Designers are free to implement circuitry within the PCS and PMA in an application-dependent manner provided that the MDI and XGMII (if the XGMII is implemented) specifications are met. When the PHY is incorporated within the physical bounds of a single-port device or a multiport device, implementation of the XGMII is optional. System operation from the perspective of signals at the MDI and management objects are identical whether the XGMII is implemented or not.

55.1.6 Conventions in this clause

The body of this clause contains state diagrams, including definitions of variables, constants, and functions. Should there be a discrepancy between a state diagram and descriptive text, the state diagram prevails.

The notation used in the state diagrams follows the conventions of 21.5.

The values of all components in test circuits shall be accurate to within −1% unless otherwise stated.

Default initializations, unless specifically specified, are left to the implementor.

55.2 10GBASE-T Service Primitives and Interfaces

10GBASE-T transfers data and control information across the following four service interfaces:

a) Ten Gigabit Media Independent Interface (XGMII)
b) Management Function Interface
c) PMA Service Interface
d) Medium Dependent Interface (MDI)

The XGMII is specified in Clause 46; the Management Function Interface is specified in Clause 28. The PMA Service Interface is defined in 55.2.2 and the MDI is defined in 55.8.

55.2.1 Management Function Interface

10GBASE-T uses the following service primitives to exchange status indications and control signals across the Management Function Interface as specified in Clause 28:

Editor's note: These primitives have been taken from Clause 40 (1000BASE-T). Are these valid for 10GBASE-T? If we need changes to these, please provide comments.

PMA_LINK.request (link_control)

PMA_LINK.indicate (link_status)

55.2.1.1 PMA_LINK.request

This primitive allows the Auto-Negotiation algorithm to enable and disable operation of the PMA as specified in 28.2.6.2.

55.2.1.1.1 Semantics of the primitive

PMA_LINK.request (link_control)

The link_control parameter can take on one of three values: SCAN_FOR_CARRIER, DISABLE, or ENABLE.

SCAN_FOR_CARRIER	Used by the Auto-Negotiation algorithm prior to receiving any fast link pulses. During this mode the PMA reports link_status=FAIL.PHY processes are disabled.
DISABLE	Set by the Auto-Negotiation algorithm in the event fast link pulses are detected. PHY processes are disabled. This allows the Auto-Negotiation algorithm to determine how to configure the link.
ENABLE	Used by Auto-Negotiation to turn control over to the PHY for data processing functions.

55.2.1.1.2 When generated

Auto-Negotiation generates this primitive to indicate a change in link_control as described in Clause 28.

55.2.1.1.3 Effect of receipt

This primitive affects operation of the PMA Link Monitor function as defined in 55.4.2.5.

55.2.1.2 PMA_LINK.indicate

This primitive is generated by the PMA to indicate the status of the underlying medium as specified in 28.2.6.1. This primitive informs the PCS, PMA PHY Control function, and the Auto-Negotiation algorithm about the status of the underlying link.

55.2.1.2.1 Semantics of the primitive

PMA_LINK.indicate (link_status)

The link_status parameter can take on one of three values: FAIL, READY, or OK.

FAIL	No valid link established.
READY(PB)	The Link Monitor function indicates that a 10GBASE-T link is intact and ready to be established.
OK(PB)	The Link Monitor function indicates that a valid 10GBASE-T link is established. Reliable reception of signals transmitted from the remote PHY is possible.

PB is a parameter that can take any value from 1 to TBD and indicates the power backoff mode.

55.2.1.2.2 When generated

The PMA generates this primitive continuously to indicate the value of link_status in compliance with the state diagram given in Figure 55–20.

55.2.1.2.3 Effect of receipt

The effect of receipt of this primitive is TBD.

55.2.2 PMA Service Interface

10GBASE-T uses the following service primitives to exchange symbol vectors, status indications, and control signals across the service interfaces:

Editor's note: These primitives have been taken from Clause 40. Are these valid for 10GBASE-T? If we need changes to these, please provide comments.

PMA_TXMODE.indicate (tx_mode)

PMA_CONFIG.indicate (config)

PMA_UNITDATA.request (tx_symb_vector)

PMA_UNITDATA.indicate (rx_symb_vector)

PMA_SCRSTATUS.request (scr_status)

PMA_RXSTATUS.indicate (loc_rcvr_status)

PMA_REMRXSTATUS.request (rem_rcvr_status)

The use of these primitives is illustrated in Figure 55–4.

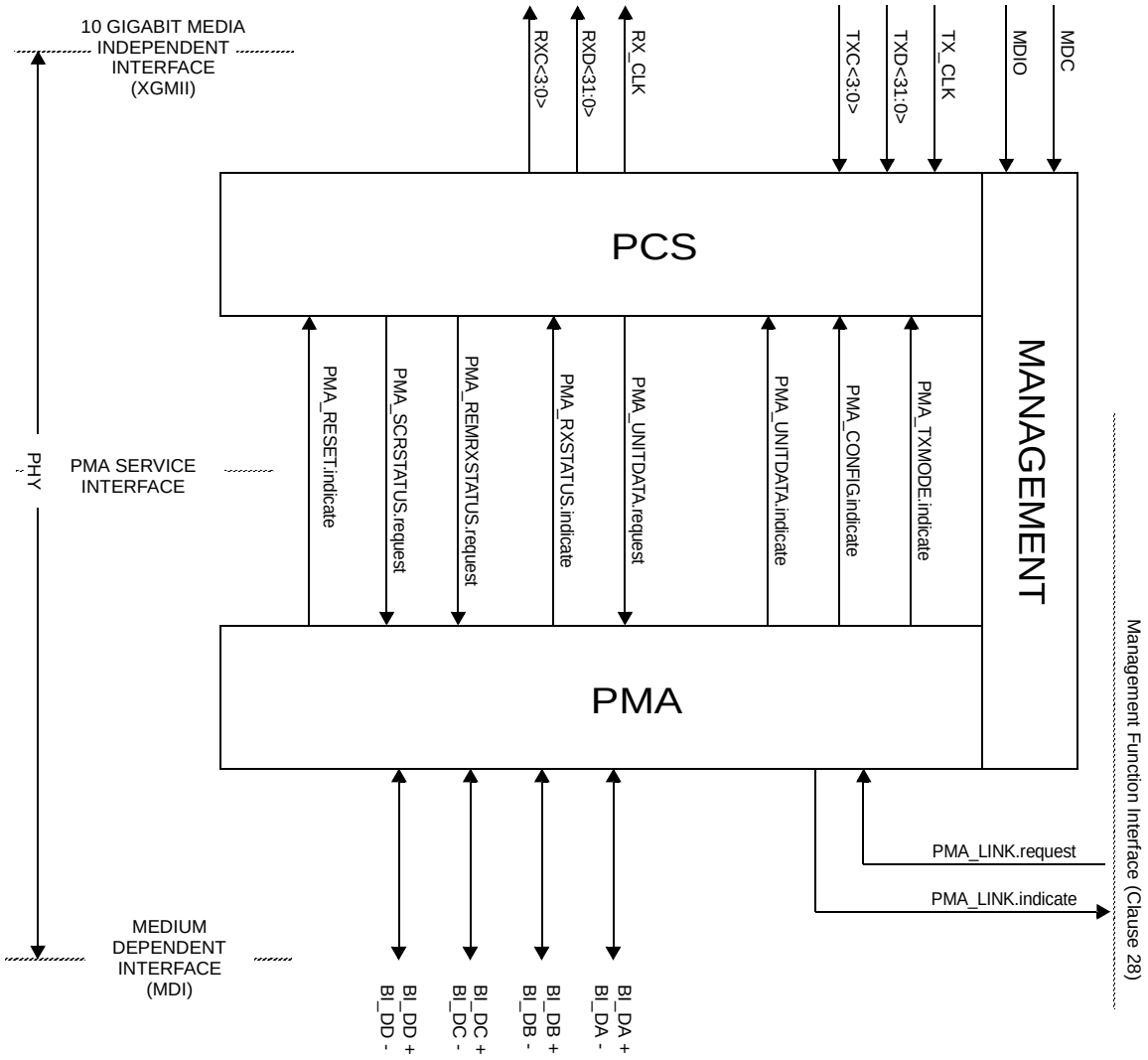


Figure 55–4—10GBASE-T service interfaces

55.2.3 PMA_TXMODE.indicate

The transmitter in a 10GBASE-T link normally sends over the four pairs, four dimensional symbols that represent an XGMII data stream with framing, scrambling and encoding of data, control information, or idles.

55.2.3.1 Semantics of the primitive

PMA_TXMODE.indicate (tx_mode)

PMA_TXMODE.indicate specifies to PCS Transmit via the parameter tx_mode what sequence of code-groups the PCS should be transmitting. The parameter tx_mode can take on one of the following three values of the form:

SEND_N This value is continuously asserted when transmission of sequences of four dimensional symbols representing an XGMII data stream in normal mode. or idle mode is to take place.

SEND_T This value is continuously asserted in case transmission of sequences of code-groups representing the startup mode is to take place.

SEND_Z This value is continuously asserted in case transmission of zeros is required.

55.2.3.2 When generated

The PMA PHY Control function generates PMA_TXMODE.indicate messages continuously.

55.2.3.3 Effect of receipt

Upon receipt of this primitive, the PCS performs its Transmit function as described in 55.3.2.2.

55.2.4 PMA_CONFIG.indicate

Each PHY in a 10GBASE-T link is capable of operating as a MASTER PHY and as a SLAVE PHY. MASTER-SLAVE configuration is determined during Auto-Negotiation (55.6.1). The result of this negotiation is provided to the PMA.

55.2.4.1 Semantics of the primitive

PMA_CONFIG.indicate (config)

PMA_CONFIG.indicate specifies to PCS and PMA Transmit via the parameter config whether the PHY must operate as a MASTER PHY or as a SLAVE PHY and the power backoff level at which the transmitter shall operate. The parameter config can take on one of the following two values of the form:

MASTER(PB) This value is continuously asserted when the PHY must operate as a MASTER PHY.

SLAVE(PB) This value is continuously asserted when the PHY must operate as a SLAVE PHY.

PB can take any value from 1 to TBD to indicate the power backoff mode.

55.2.4.2 When generated

PMA generates PMA_CONFIG.indicate messages continuously.

55.2.4.3 Effect of receipt

PCS and PMA Clock Recovery perform their functions in MASTER or SLAVE configuration according to the value assumed by the parameter config.

55.2.5 PMA_UNITDATA.request

This primitive defines the transfer of code-groups in the form of the tx_symb_vector parameter from the PCS to the PMA. The code-groups are obtained in the PCS Transmit function using the encoding rules defined in 55.3.2.2 to represent XGMII data and control streams or other sequences.

55.2.5.1 Semantics of the primitive

PMA_UNITDATA.request (tx_symb_vector)

During transmission, the PMA_UNITDATA.request simultaneously conveys to the PMA via the parameter tx_symb_vector the value of the symbols to be sent over each of the four transmit pairs BI_DA, BI_DB, BI_DC, and BI_DD. The tx_symb_vector parameter takes on the form:

SYMB_4D — A vector of four multi-level symbols, one for each of the four transmit pairs BI_DA, BI_DB, BI_DC, and BI_DD. Each symbol may take on one of the values in the set {-TBD,...,+TBD}.

The symbols that are elements of tx_symb_vector are called, according to the pair on which each will be transmitted, tx_symb_vector[BI_DA], tx_symb_vector[BI_DB], tx_symb_vector[BI_DC], and tx_symb_vector[BI_DD].

55.2.5.2 When generated

The PCS generates PMA_UNITDATA.request (SYMB_4D) synchronously with every transmit clock cycle.

55.2.5.3 Effect of receipt

Upon receipt of this primitive the PMA transmits on the MDI the signals corresponding to the indicated symbols after processing with the Tomlinson Harashima Precoder (THP), the transmit filter and other specified PMA transmit processing. The parameter tx_symb_vector is also used by the PMA Receive function to process the signals received on pairs BI_DA, BI_DB, BI_DC, and BI_DD for cancelling the echo and Near End Cross Talk (NEXT).

55.2.6 PMA_UNITDATA.indicate

This primitive defines the transfer of code-groups in the form of the rx_symb_vector parameter from the PMA to the PCS.

55.2.6.1 Semantics of the primitive

PMA_UNITDATA.indicate (rx_symb_vector)

During reception the PMA_UNITDATA.indicate simultaneously conveys to the PCS via the parameter rx_symb_vector the values of the symbols detected on each of the four receive pairs BI_DA, BI_DB, BI_DC, and BI_DD. The rx_symb_vector parameter takes on the form:

SYMB_4D — A vector of 4 symbols that is the receivers best estimate of the information bits that were sent by the remote transmitter.

55.2.6.2 When generated

The PMA generates PMA_UNITDATA.indicate (SYMB_4D) messages synchronously every 4 symbols received at the MDI. The nominal rate of the PMA_UNITDATA.indicate primitive is TBD MHz, as governed by the recovered clock.

55.2.6.3 Effect of receipt

The effect of receipt of this primitive is unspecified.

55.2.7 PMA_SCRSTATUS.request

This primitive is generated by PCS Receive to communicate the status of the descrambler for the local PHY. The parameter scr_status conveys to the PMA Receive function the information that the descrambler has achieved synchronization.

Editor's note: Do we need to create a similar primitive to communicate status of PCS descrambler?

55.2.7.1 Semantics of the primitive

PMA_SCRSTATUS.request (scr_status)

The scr_status parameter can take on one of two values of the form:

OK	The descrambler has achieved synchronization.
NOT_OK	The descrambler is not synchronized.

55.2.7.2 When generated

PCS Receive generates PMA_SCRSTATUS.request messages continuously.

55.2.7.3 Effect of receipt

The effect of receipt of this primitive is specified in 55.4.2.3, 55.4.2.4, and 55.4.6.1.

55.2.8 PMA_RXSTATUS.indicate

This primitive is generated by PMA Receive to indicate the status of the receive link at the local PHY. The parameter loc_rcvr_status conveys to the PCS Transmit, PCS Receive, PMA PHY Control function, and Link Monitor the information on whether the status of the overall receive link is satisfactory or not. Note that loc_rcvr_status is used by the PCS Receive decoding functions. The criterion for setting the parameter loc_rcvr_status is left to the implementor. It can be based, for example, on observing the mean-square error at the decision point of the receiver and detecting errors during reception of symbol streams.

55.2.8.1 Semantics of the primitive

PMA_RXSTATUS.indicate (loc_rcvr_status)

The loc_rcvr_status parameter can take on one of two values of the form:

OK	This value is asserted and remains true during reliable operation of the receive link for the local PHY.
NOT_OK	This value is asserted whenever operation of the link for the local PHY is unreliable.

55.2.8.2 When generated

PMA Receive generates PMA_RXSTATUS.indicate messages continuously on the basis of signals received at the MDI.

55.2.8.3 Effect of receipt

The effect of receipt of this primitive is specified in Figure 55–19 and in subclauses 55.2 and 55.4.6.2.

55.2.9 PMA_REMRXSTATUS.request

This primitive is generated by PCS Receive to indicate the status of the receive link at the remote PHY as communicated by the remote PHY via its encoding of its loc_rcvr_status parameter. The parameter rem_rcvr_status conveys to the PMA PHY Control function the information on whether reliable operation of the remote PHY is detected or not. The criterion for setting the parameter rem_rcvr_status is left to the implementor. It can be based, for example, on asserting rem_rcvr_status is NOT_OK until loc_rcvr_status is OK and then asserting the detected value of rem_rcvr_status after proper PCS receive decoding is achieved.

55.2.9.1 Semantics of the primitive

PMA_REMRXSTATUS.request (rem_rcvr_status)

The rem_rcvr_status parameter can take on one of two values of the form:

OK	The receive link for the remote PHY is operating reliably.
NOT_OK	Reliable operation of the receive link for the remote PHY is not detected.

55.2.9.2 When generated

The PCS generates PMA_REMRXSTATUS.request messages continuously on the basis on signals received at the MDI.

55.2.9.3 Effect of receipt

The effect of receipt of this primitive is specified in Figure 55–19.

55.2.10 PMA_RESET.indicate

This primitive is used to pass the PMA Reset function to the PCS (pcs_reset=ON) when reset is enabled.

The PMA_RESET.indicate primitive can take on one of two values:

TRUE	Reset is enabled.
FALSE	Reset is not enabled.

55.2.10.1 When generated

The PMA Reset function is executed as described in 55.4.2.1.

55.2.10.2 Effect of receipt

The effect of receipt of this primitive is specified in 55.4.2.1.

55.3 Physical Coding Sublayer (PCS)

Editor's note: Since both the PAM12 and PAM8 proposals as based on 64B/65B blocks, Clause 49 (PCS for 64/66B, type 10GBASE-R) has been used as the starting point for most of the PCS for this initial draft of 10GBASE-T in addition to some sections of Clause 40 (1000BASE-T). Please provide comments.

55.3.1 PCS service interface (XGMII)

The PCS service interface allows the 10GBASE-T PCS to transfer information to and from a PCS client. The PCS Interface is precisely defined as the 10 Gigabit Media Independent Interface (XGMII) in Clause 46.

55.3.2 PCS functions

The PCS comprises one PCS Reset function and two simultaneous and asynchronous operating functions. The PCS operating functions are: PCS Transmit and PCS Receive. All operating functions start immediately after the successful completion of the PCS Reset function.

The PCS reference diagram, Figure 55–5, shows how the two operating functions relate to the messages of the PCS-PMA interface. Connections from the management interface (signals MDC and MDIO) to other layers are pervasive, and are not shown in Figure 55–5. Management is specified in Clause 30.

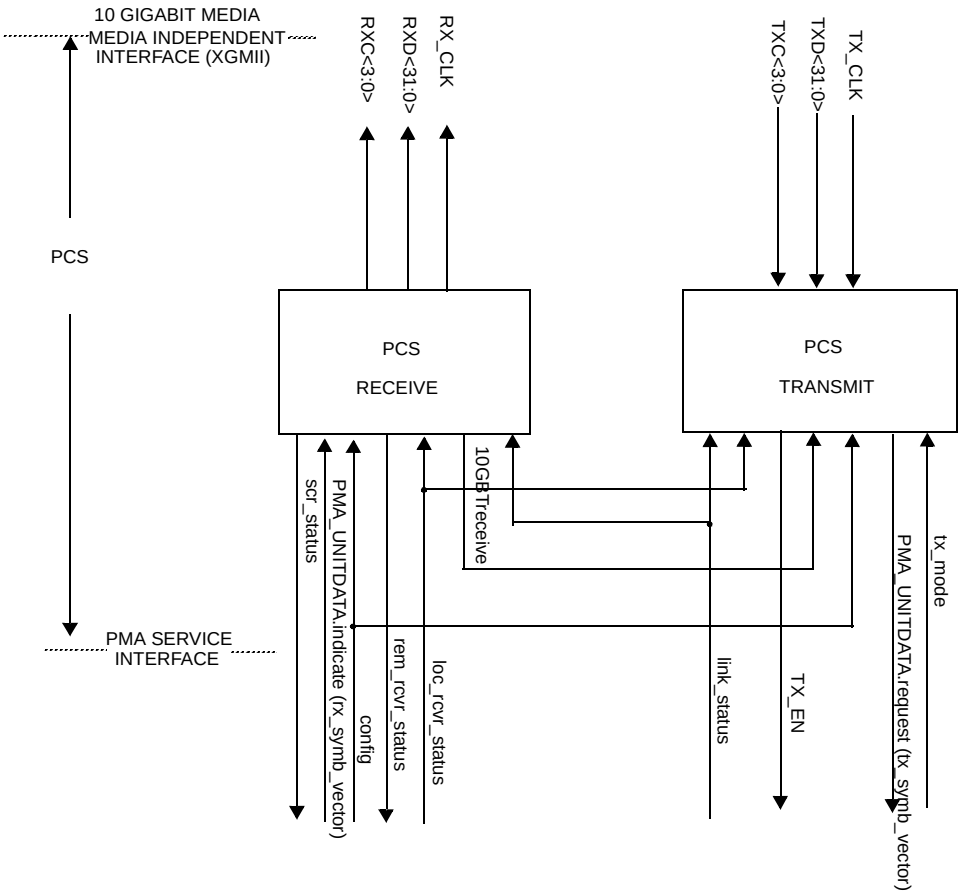


Figure 55–5—PCS reference diagram

55.3.2.1 PCS Reset function

PCS Reset initializes all PCS functions. The PCS Reset function shall be executed whenever one of the following conditions occur:

a) Power on (see 36.2.5.1.3).
b) The receipt of a request for reset from the management entity.

PCS Reset sets pcs_reset=ON while any of the above reset conditions hold true. All state diagrams take the open-ended pcs_reset branch upon execution of PCS Reset. The reference diagrams do not explicitly show the PCS Reset function.

55.3.2.2 PCS Transmit function

The PCS Transmit function shall conform to the PCS Transmit state diagram in Figure 55–15.

When communicating with the XGMII, the PCS uses a four octet-wide, synchronous data path, with packet delimiting being provided by transmit control signals and receive control signals. Alignment to 64B/65B block is performed in the PCS. The PMA sublayer operate independent of block and packet boundaries. The PCS provides the functions necessary to map packets between the XGMII format and the PMA service interface format.

When the transmit channel is in normal mode, the PCS Transmit process continuously generates 65B blocks based upon the TXD <31:0> and TXC <3:0> signals on the XGMII. The subsequent functions of the PCS Transmit process then pack the resulting bits into Low Density Parity Check (LDPC) blocks and LDPC frames (TBD). Transmit data-units are sent to the PMA or service interface via the PMA_UNITDATA.request primitive, respectively.

When the receive channel is in normal mode, the PCS Synchronization process continuously monitors PMA_SIGNAL.indicate(SIGNAL_OK). When SIGNAL_OK indicates OK, then the PCS Synchronization process accepts data-units via the PMA_UNITDATA.indicate primitive. It attains frame and block synchronization based on the LDPC frame (TBD) synchronization headers and conveys received blocks to the PCS Receive process. The PCS Synchronization process sets the sync_status flag to indicate whether the PCS has obtained synchronization.

When the PCS Synchronization process has obtained synchronization, the LDPC Block Error Rate (LBER) monitor process monitors the signal quality asserting hi_lber if excessive errors are detected. When sync_status is asserted and hi_lber is de-asserted, the PCS Receive process continuously accepts blocks. The PCS Receive process monitors these blocks and generates RXD <31:0> and RXC <3:0> on the XGMII.

In each symbol period, when communicating with the PMA, the PCS Transmit generates a code-group (A_n, B_n, C_n, D_n) that is transferred to the PMA via the PMA_UNITDATA.request primitive. The PMA transmits symbols A_n, B_n, C_n, D_n over wire-pairs BI_DA, BI_DB, BI_DC, and BI_DD respectively. The integer, n, is a time index that is introduced to establish a temporal relationship between different symbol periods. A symbol period, T, is nominally equal to TBD ns. If a PMA_TXMODE.indicate message has the value SEND_Z, PCS Transmit passes a vector of zeros at each symbol period to the PMA via the PMA_UNITDATA.request primitive.

Editor's note: Clause 49 transports idle as 64B/65B control payload. Should we preserve a PMA idle mode with binary level PAM for PMA training?

If a PMA_TXMODE.indicate message has the value SEND_T, PCS Transmit generates sequences of code-groups according to the encoding rule in training mode. Special code-groups that use only the values {-L, L} (L TBD) are transmitted in this case. Training mode encoding also takes into account the value of the parameter loc_rcvr_status. By this mechanism, a PHY indicates the status of its own receiver to the link partner.

In the normal mode of operation, the PMA_TXMODE.indicate message has the value SEND_N, and the PCS Transmit function uses a 65B-LDPC coding technique to generate at each symbol period code-groups that represent data or control. During transmission, the 65B bits are scrambled by the PCS using a PCS scrambler, then encoded into a code-group of four dimensional symbols and transferred to the PMA. During data encoding, PCS Transmit utilizes a block LDPC encoder.

55.3.3 Use of blocks

The PCS maps XGMII signals into 65-bit blocks inserted into frames of LDPC blocks, and vice versa, using a 65B-LDPC coding scheme. The LDPC frame synchronization headers of the blocks allow establishment of block boundaries by the PCS Synchronization process. Blocks are unobservable and have no meaning outside the PCS. The PCS functions ENCODE and DECODE generate, manipulate, and interpret blocks as provided by the rules in 55.3.4.

55.3.4 65B-LDPC transmission code

The PCS uses a transmission code to improve the transmission characteristics of information to be transferred across the link and to support transmission of control and data characters. The encoding defined by the transmission code ensure that sufficient information is present in the PHY bit stream to make clock recovery possible at the receiver. The encoding also preserves the likelihood of detecting any LDPC block errors that may occur during transmission and reception of information. In addition, the synchronization headers of the code enable the receiver to achieve block alignment on the incoming PHY bit stream.

The relationship of block bit positions to XGMII, PMA, and other PCS constructs is illustrated in Figure 55–6 for transmit and Figure 55–7 for receive. These figures illustrate the processing of a multiplicity of blocks containing 8 data octets. See 55.3.4.3 for information on how blocks containing control characters are mapped.

55.3.4.1 Notation conventions

For values shown as binary, the leftmost bit is the first transmitted bit.

64B/65B encodes 8 data octets or control characters into a block. Blocks containing control characters also contain a block type field. Data octets are labeled D_0 to D_7. Control characters other than /O/, /S/ and /T/ are labeled C_0 to C_7. The control character for ordered_set is labeled as O_0 or O_4 since it is only valid on the first octet of the XGMII. The control character for start is labeled as S_0 or S_4 for the same reason. The control character for terminate is labeled as T_0 to T_7.

Two consecutive XGMII transfers provide eight characters that are encoded into one 65-bit transmission block. The subscript in the above labels indicates the position of the character in the eight characters from the XGMII transfers.

Contents of block type fields, data octets and control characters are shown as hexadecimal values. The LSB of the hexadecimal value represents the first transmitted bit. For instance, the block type field 0x1e is sent from left to right as 01111000. The bits of a transmitted or received block are labeled TxB<64:0> and RxB<64:0> respectively where TxB<0> and RxB<0> represent the first transmitted bit. The value of the data/ctrl header is shown as a binary value. Binary values are shown with the first transmitted bit (the LSB) on the left.

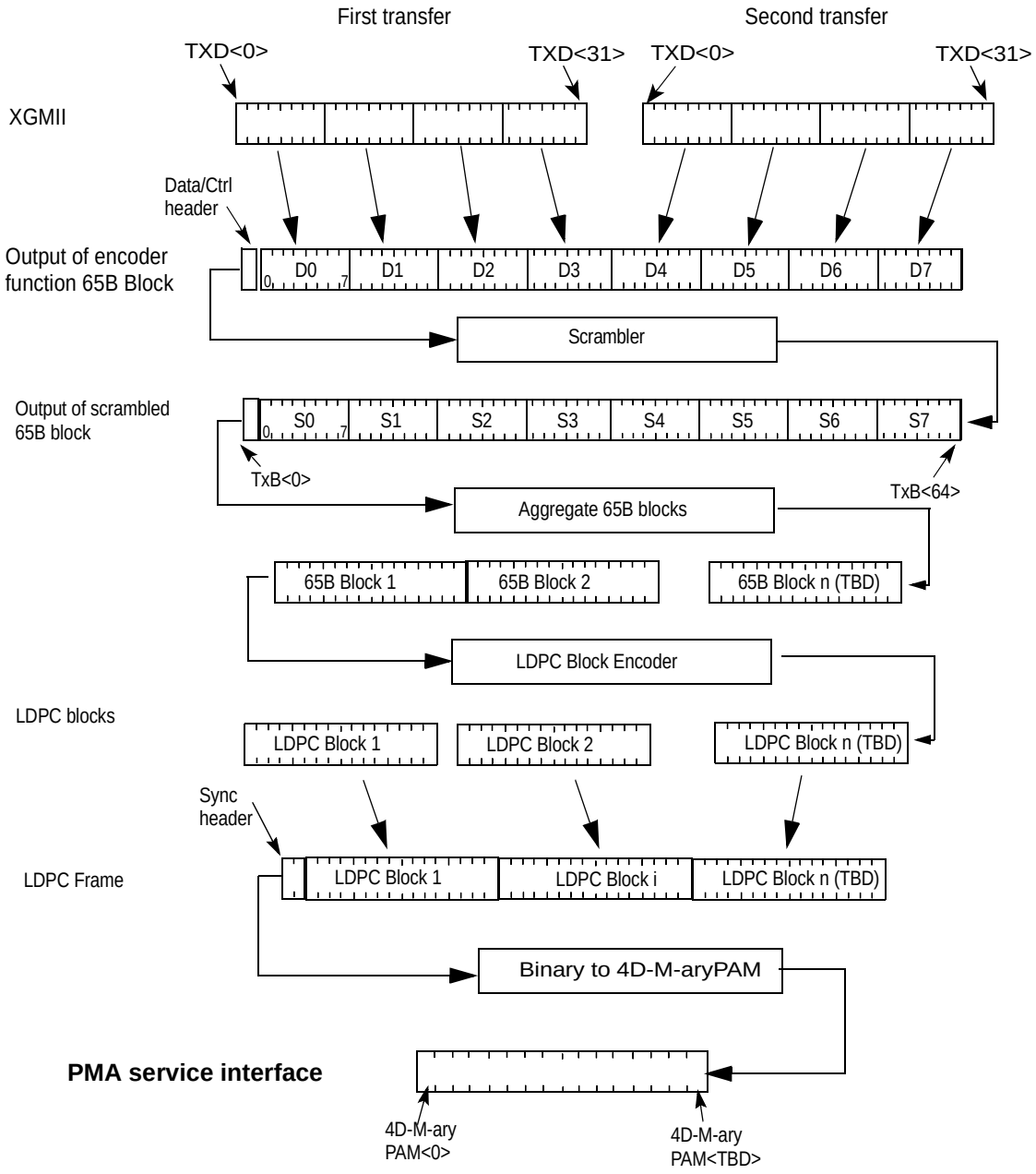


Figure 55–6—PCS Transmit Bit Ordering

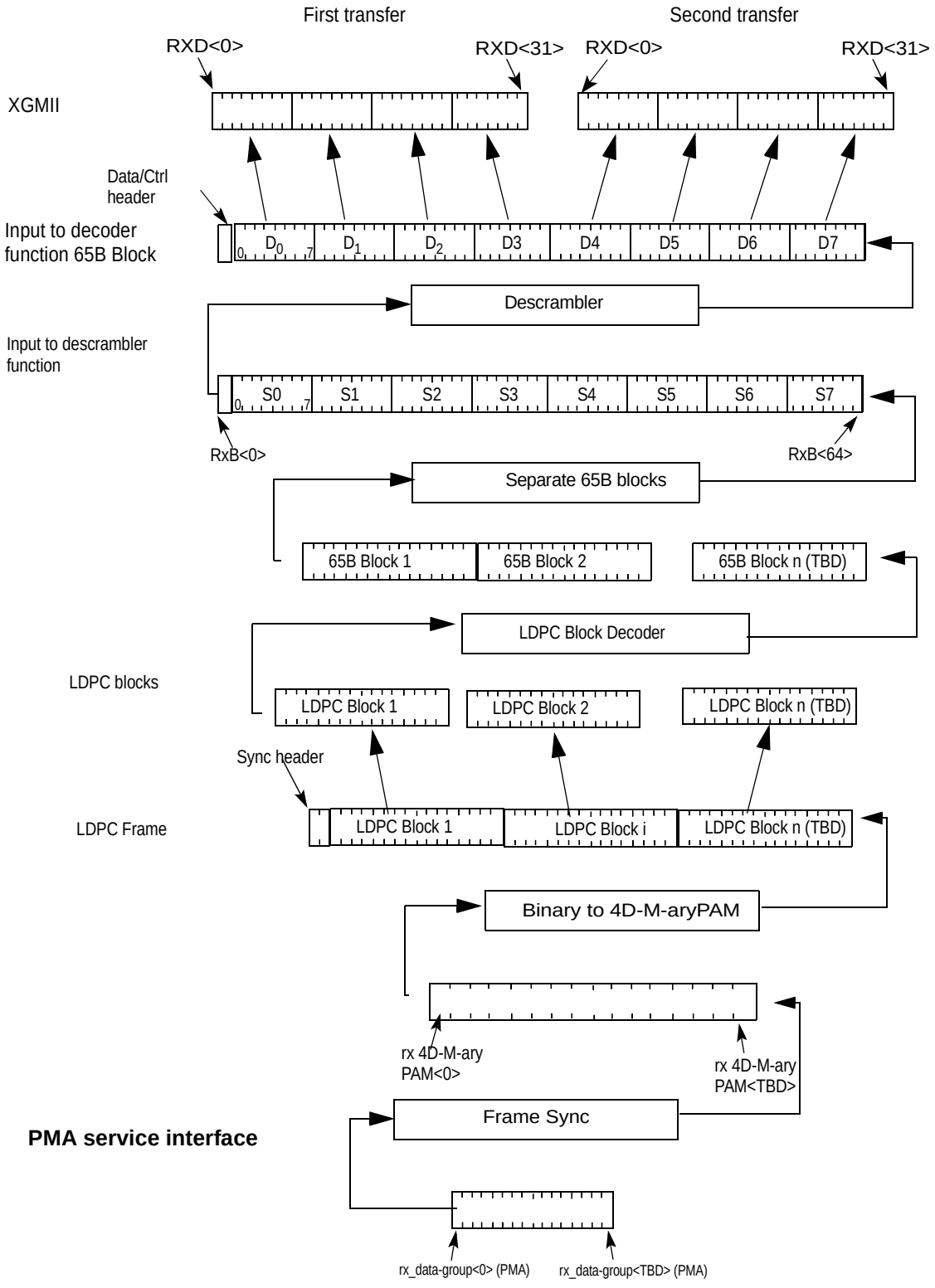


Figure 55–7—PCS Receive Bit Ordering

55.3.4.2 Transmission order

Block bit transmission order is illustrated in Figure 55–6 and Figure 55–7. Note that these figures show the mapping from XGMII to 64B/65B block for a block containing eight data characters.

55.3.4.3 Block structure

Blocks consist of 65 bits. The first bit of a block is the data/ctrl header. Blocks are either data blocks or control blocks. The data/ctrl header is 0 for data blocks and 1 for control blocks. The remainder of the block contains the payload.

Data blocks contain eight data characters. Control blocks begin with an 8-bit block type field that indicates the format of the remainder of the block. For control blocks containing a Start or Terminate character, that character is implied by the block type field. Other control characters are encoded in a 7-bit control code or a 4-bit O Code. Each control block contains eight characters.

The format of the blocks is as shown in Figure 55–8. In the figure, the column labeled Input Data shows, in abbreviated form, the eight characters used to create the 65-bit block. These characters are either data characters or control characters and, when transferred across the XGMII interface, the corresponding TXC or RXC bit is set accordingly. Within the Input Data column, D_0 through D_7 are data octets and are transferred with the corresponding TXC or RXC bit set to zero. All other characters are control octets and are transferred with the corresponding TXC or RXC bit set to one. The single bit fields (thin rectangles with no label in the figure) are sent as zero and ignored upon receipt.

Bits and field positions are shown with the least significant bit on the left. Hexadecimal numbers are shown in normal hexadecimal. For example the block type field 0x1e is sent as 01111000 representing bits 1 through 8 of the 65 bit block. The least significant bit for each field is placed in the lowest numbered position of the field.

All unused values of block type field[1] are reserved.

55.3.4.4 Control codes

The same set of control characters are supported by the XGMII and the 10GBASE-T PCS. The representations of the control characters are the control codes. XGMII encodes a control character into an octet (an eight bit value). The 10GBASE-T PCS encodes the start and terminate control characters implicitly by the block type field. The 10GBASE-T PCS encodes the ordered_set control codes using a combination of the block type field and a 4-bit O code for each ordered_set. The 10GBASE-T PCS encodes each of the other control characters into a 7-bit C code).

The control characters and their mappings to 10GBASE-T control codes and XGMII control codes are specified in Table 55–1. All XGMII and 10GBASE-T control code values that do not appear in the table shall not be transmitted and shall be treated as an error if received.

55.3.4.5 Ordered sets

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. Ordered sets consist of a control character followed by three data characters. Ordered sets always begin on the first octet of the XGMII. 10 Gigabit Ethernet uses one kind of ordered_set: the sequence ordered_set (see 46.3.4). The sequence ordered_set control character is denoted /Q/. An addi-

[1]The block type field values have been chosen to have a 4-bit Hamming distance between them. The only unused value which maintains the Hamming distance is 0x00.

Input Data	data ctrl bit	Block Payload							
Bit Position: / Data Block Format:	0	1							64
D_0 D_1 D_2 D_3/D_4 D_5 D_6 D_7	0	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7
Control Block Formats:		**Block Type Field**							
C_0 C_1 C_2 C_3/C_4 C_5 C_6 C_7	1	0x1e	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7
C_0 C_1 C_2 C_3/O_4 D_5 D_6 D_7	1	0x2d	C_0	C_1	C_2	C_3	O_4	D_5	D_6	D_7
C_0 C_1 C_2 C_3/S_4 D_5 D_6 D_7	1	0x33	C_0	C_1	C_2	C_3		D_5	D_6	D_7
O_0 D_1 D_2 D_3/S_4 D_5 D_6 D_7	1	0x66	D_1	D_2	D_3	O_0		D_5	D_6	D_7
O_0 D_1 D_2 D_3/O_4 D_5 D_6 D_7	1	0x55	D_1	D_2	D_3	O_0	O_4	D_5	D_6	D_7
S_0 D_1 D_2 D_3/D_4 D_5 D_6 D_7	1	0x78	D_1	D_2	D_3	D_4		D_5	D_6	D_7
O_0 D_1 D_2 D_3/C_4 C_5 C_6 C_7	1	0x4b	D_1	D_2	D_3	O_0	C_4	C_5	C_6	C_7
T_0 C_1 C_2 C_3/C_4 C_5 C_6 C_7	1	0x87		C_1	C_2	C_3	C_4	C_5	C_6	C_7
D_0 T_1 C_2 C_3/C_4 C_5 C_6 C_7	1	0x99	D_0		C_2	C_3	C_4	C_5	C_6	C_7
D_0 D_1 T_2 C_3/C_4 C_5 C_6 C_7	1	0xaa	D_0	D_1		C_3	C_4	C_5	C_6	C_7
D_0 D_1 D_2 T_3/C_4 C_5 C_6 C_7	1	0xb4	D_0	D_1	D_2		C_4	C_5	C_6	C_7
D_0 D_1 D_2 D_3/T_4 C_5 C_6 C_7	1	0xcc	D_0	D_1	D_2	D_3		C_5	C_6	C_7
D_0 D_1 D_2 D_3/D_4 T_5 C_6 C_7	1	0xd2	D_0	D_1	D_2	D_3	D_4		C_6	C_7
D_0 D_1 D_2 D_3/D_4 D_5 T_6 C_7	1	0xe1	D_0	D_1	D_2	D_3	D_4	D_5		C_7
D_0 D_1 D_2 D_3/D_4 D_5 D_6 T_7	1	0xff	D_0	D_1	D_2	D_3	D_4	D_5	D_6	

Figure 55–8—64B/65B Block Formats

tional ordered_set, the signal ordered_set, has been reserved and it begins with another control code. The 4-bit O field encodes the control code. See Table 55–1 for the mappings.

55.3.4.6 Valid and invalid blocks

A block is invalid if any of the following conditions exists:

a) The block type field contains a reserved value.
b) Any control character contains a value not in Table 55–1.
c) Any O code contains a value not in Table 55–1.
d) The set of eight XGMII characters does not have a corresponding block format in Figure 55–8.

55.3.4.7 Idle (/I/)

Idle control characters (/I/) are transmitted when idle control characters are received from the XGMII. Idle characters may be added or deleted by the PCS to adapt between clock rates. /I/ insertion and deletion shall occur in groups of 4. /I/s may be added following idle or ordered sets. They shall not be added while data is being received. When deleting /I/s, the first four characters after a /T/ shall not be deleted.

Table 55–1—Control Codes

Control Character	Notation	XGMII Control Code	10GBASE-T Control Code	10GBASE-T O Code	8B/10B Code[a]
idle	/I/	0x07	0x00		K28.0 or K28.3 or K28.5
start	/S/	0xfb	Encoded by block type field		K27.7
terminate	/T/	0xfd	Encoded by block type field		K29.7
error	/E/	0xfe	0x1e		K30.7
Sequence ordered_set	/Q/	0x9c	Encoded by block type field plus O code	0x0	K28.4
reserved0	/R/[b]	0x1c	0x2d		K28.0
reserved1		0x3c	0x33		K28.1
reserved2	/A/	0x7c	0x4b		K28.3
reserved3	/K/	0xbc	0x55		K28.5
reserved4		0xdc	0x66		K28.6
reserved5		0xf7	0x78		K23.7
Signal ordered_set[c]	/Fsig/	0x5c	Encoded by block type field plus O code	0xF	K28.2

[a]For information only. The 8B/10B code is specified in Clause 36. Usage of the 8B/10B code for 10 Gb/s operation is specified in Clause 48.
[b]The codes for /A/, /K/, and /R/ are used on the XAUI interface to signal idle. They are not present on the XGMII when no errors have occurred, but certain bit errors cause the XGXS to send them on the XGMII.
[c]Reserved for INCITS T11 Fibre Channel use.

55.3.4.8 Start (/S/)

The start control character (/S/) indicates the start of a packet. This delimiter is only valid on the first octet of the XGMII (TXD<0:7> and RXD<0:7>). Receipt of an /S/ on any other octet of TxD indicates an error. Block type field values implicitly encode an /S/ as the fifth or first character of the block. These are the only characters of a block on which a start can occur.

55.3.4.9 Terminate (/T/)

The terminate control character (/T/) indicates the end of a packet. Since packets may be any length, the /T/ can occur on any octet of the XGMII interface and within any character of the block. The location of the /T/ in the block is implicitly encoded in the block type field. A valid end of packet occurs when a block containing a /T/ is followed by a control block that does not contain a /T/.

55.3.4.10 ordered_set (/O/)

The ordered_set control characters (/O/) indicate the start of an ordered_set. There are two kinds of ordered sets: the sequence ordered_set and the signal ordered_set (which is reserved). When it is necessary to designate the control character for the sequence ordered_set specifically, /Q/ will be used. /O/ is only valid on the first octet of the XGMII. Receipt of an /O/ on any other octet of TXD indicates an error. Block type field values implicitly encode an /O/ as the first or fifth character of the block. The 4-bit O code encodes the specific /O/ character for the ordered_set.

Sequence ordered_sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and shall delete only one of the two. Only Idles may be inserted for clock compensation. Signal ordered_sets are not deleted for clock compensation.

55.3.4.11 Error (/E/)

The /E/ is sent whenever an /E/ is received. It is also sent when invalid blocks are received. The /E/ allows physical sublayers such as the XGXS and PCS to propagate received errors. See R_BLOCK_TYPE and T_BLOCK_TYPE function definitions in 55.3.12.2.3 for further information.

55.3.5 Transmit process

The transmit process generates blocks based upon the TXD<31:0> and TXC<3:0> signals received from the XGMII. Two XGMII data transfers are encoded into each block. It takes TBD PMA_UNITDATA transfers to send a block of data. Therefore, if the PCS is connected to an XGMII and PMA sublayer where the ratio of their transfer rates is exactly TBD1:TBD2, then the transmit process does not need to perform rate adaptation. Where the XGMII and PMA sublayer data rates are not synchronized to that ratio, the transmit process will need to insert idles, delete idles, or delete sequence ordered sets to adapt between the rates.

The transmit process generates blocks as specified in the transmit process state machine. The contents of each block are contained in a vector tx_coded<64:0>, which is passed to the scrambler. tx_coded<0> contains the data/ctrl header and the remainder of the bits contain the block payload.

55.3.6 PCS Scrambler.

The payload of the block is scrambled with a self-synchronizing scrambler. The scrambler shall produce the same result as the implementation shown in Figure 55–9. This implements the scrambler polynomial:[2]

$$G(x) = 1 + x^{39} + x^{58} \quad (55–1)$$

There is no requirement on the initial value for the scrambler. In no case shall the scrambler state be initialized to all zeros. The scrambler is run continuously on all payload bits.

55.3.7 65B-LDPC

The 65B/LDPC adapts between the 65-bit width of the 65B blocks and the 4D-M-aryPAM width of the PMA. When the transmit channel is operating in normal mode, the 65B-LDPC sends 4D-M-aryPAM of transmit data at a time via PMA_UNITDATA.request primitives. The UNITDATA.request primitives are fully packed with bits.

[2]The convention here, which considers the most recent bit into the scrambler to be the lowest order term, is consistent with most references and with other scramblers shown in this standard. Some references consider the most recent bit into the scrambler to be the highest order term and would therefore identify this as the inverse of the polynomial in Equation (55–1). In case of doubt, note that the conformance requirement is based on the representation of the scrambler in the figure rather than the polynomial equation.

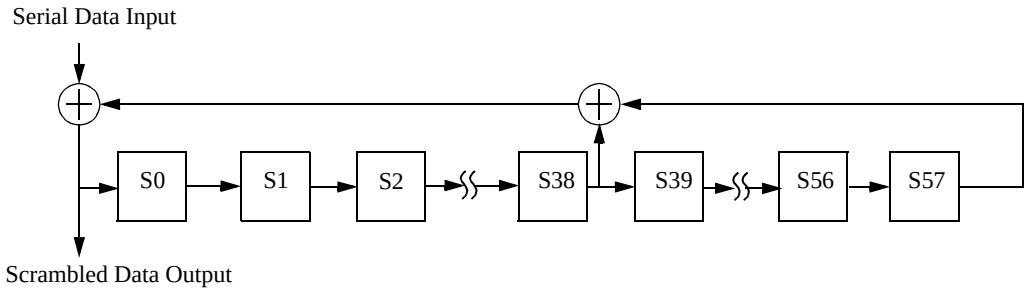


Figure 55–9—Scrambler

55.3.8 Block synchronization

When the receive channel is operating in normal mode, the block synchronization function receives data via 4D-M-aryPAM PMA_UNITDATA.request primitives. It shall form a 4D M-aryPAM stream from the primitives by concatenating requests with the PAMs of each primitive in order from rx_data-group<0> to rx_data-group<TBD> (see Figure 55–7). It obtains lock to the LDPC frames and the 65-bit blocks in the bit stream using the frame sync headers and outputs 65-bit blocks. Lock is obtained as specified in the block lock state machine shown in Figure 55–13.

55.3.9 PCS Descrambler

The descrambler processes the payload to reverse the effect of the scrambler using the same polynomial. It shall produce the same result as the implementation shown in Figure 55–10.

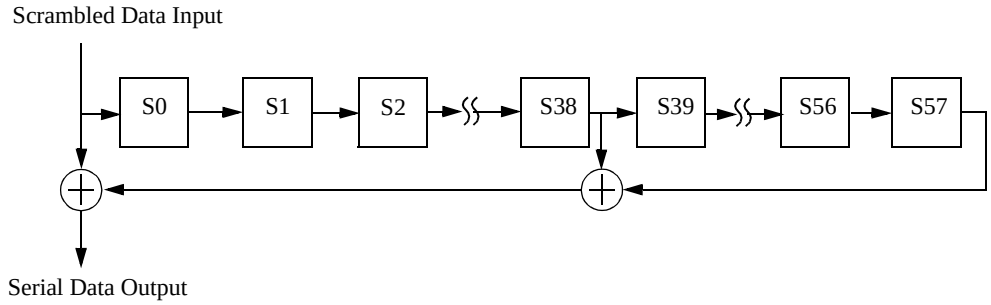


Figure 55–10—Descrambler

55.3.10 Receive function

The PCS Receive function accepts received code-groups provided by the PMA Receive function via the parameter rx_symb_vector. To achieve correct operation, PCS Receive uses the knowledge of the encoding rules that are employed. PCS Receive checks the received framing and signals the reliable acquisition of the descrambler state by setting the parameter scr_status to OK. The received 65B-LDPC frames are decoded into blocks of 65-bits to obtain the signals RXD<31:0> and RXC<3:0> for transmission to the XGMII. Two XGMII data transfers are decoded from each block. Where the XGMII and PMA sublayer data rates are not synchronized to a TBD1:TBD2 ratio, the receive process will insert idles, delete idles, or delete sequence ordered sets to adapt between rates. PCS Receive detects the transmission of a stream of data from the remote station and conveys this information to the PCS Transmit functions via the parameter 10GBTreceive.

The receive process decodes blocks as specified in the receive state machine shown in Figure 55–16.

55.3.11 PMA Side-stream scrambler polynomials

Editor's note: For PMA training the side-stream scrambler from 1000BASE-T has been used for this initial draft. This results in two different scramblers, one from Clause 49 for scrambling the 65B blocks and the second from Clause 40 for PAM training. Please provide comments.

The PCS Transmit function employs side-stream scrambling for 2-level PAM PMA training. If the parameter config provided to the PCS by the PMA PHY Control function via the PMA_CONFIG.indicate message assumes the value MASTER, PCS Transmit shall employ

$$g_M(x) = 1 + x^{13} + x^{33}$$

as transmitter side-stream scrambler generator polynomial. If the PMA_CONFIG.indicate message assumes the value of SLAVE, PCS Transmit shall employ

$$g_S(x) = 1 + x^{20} + x^{33}$$

as transmitter side-stream scrambler generator polynomial. An implementation of master and slave PHY side-stream scramblers by linear-feedback shift registers is shown in Figure 55–11. The bits stored in the shift register delay line at time n are denoted by $Scr_n[32:0]$. At each symbol period, the shift register is advanced by one bit, and one new bit represented by $Scr_n[0]$ is generated. The transmitter side-stream scrambler is reset upon execution of the PCS Reset function. If PCS Reset is executed, all bits of the TBD-bit vector representing the side-stream scrambler state are arbitrarily set. The initialization of the scrambler state is left to the implementor. In no case shall the scrambler state be initialized to all zeros.

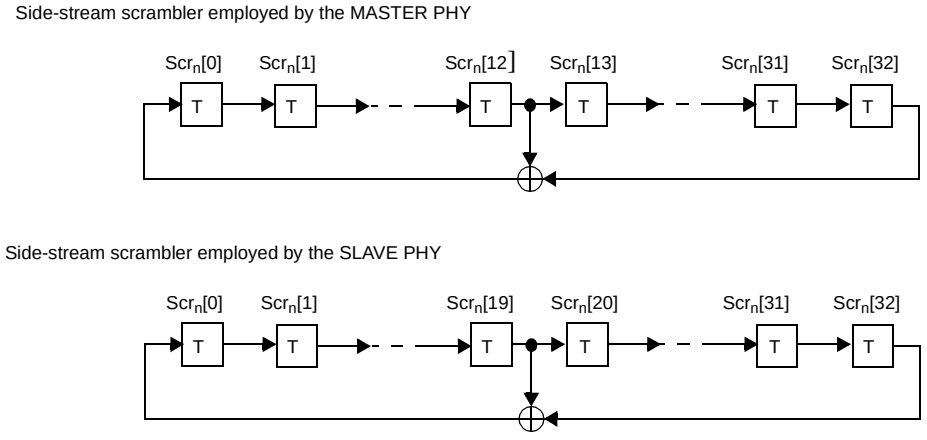


Figure 55–11—A realization of side-stream scramblers by linear feedback shift registers

55.3.11.1 Generation of bits $Sx_n[3:0]$, $Sy_n[3:0]$, and $Sg_n[3:0]$

TBD.

55.3.11.2 Generation of 4-D symbols TA_n, TB_n, TC_n, TD_n

TBD.

55.3.11.3 Decoding of code-groups

When the PMA indicates that correct receiver operation has been achieved by setting the loc_rcvr_status parameter to the value OK, the PCS Receive continuously checks that the received sequence satisfies the encoding rule. When a violation is detected, PCS Receive assigns the value TRUE to the parameter 10GBTreceive.

Upon detection of a receiver error, the signal RX_ER is asserted and the parameter rxerror_status assumes the value ERROR. De-assertion of RX_ER and transition to the IDLE state (rxerror_status=NO_ERROR) takes place upon detection of TBD consecutive vectors satisfying the encoding rule.

During reception of a stream of data, PCS Receive checks that the receive blocks. If a violation of the encoding rules is detected, PCS Receive asserts the signal TBD for at least TBD symbol periods.

A premature stream termination is caused by the detection of TBD invalid blocks during the reception of a data stream. Then, PCS Receive waits for the reception of TBD consecutive vectors satisfying the encoding rule TBD prior to de-asserting the error indication. The signal RX_ER is also asserted in the LINK FAILED state, which ensures that RX_ER remains asserted for at least one symbol period.

55.3.11.4 Receiver descrambler polynomials

The PHY shall descramble the PMA training stream and return the proper sequence of code-groups. For side-stream descrambling, the MASTER PHY shall employ the receiver descrambler generator polynomial $g'_M(x) = 1 + x^{20} + x^{33}$ and the SLAVE PHY shall employ the receiver descrambler generator polynomial $g'_S(x) = 1 + x^{13} + x^{33}$.

55.3.12 Detailed functions and state diagrams

55.3.12.1 State diagram conventions

The body of this subclause is comprised of state diagrams, including the associated definitions of variables, constants, and functions. Should there be a discrepancy between a state diagram and descriptive text, the state diagram prevails.

The notation used in the state diagrams follows the conventions of 21.5. State diagram timers follow the conventions of 14.2.3.2. The notation ++ after a counter or integer variable indicates that its value is to be incremented.

55.3.12.2 State diagram parameters

55.3.12.2.1 Constants

EBLOCK_R<71:0>
: 72 bit vector to be sent to the XGMII interface containing /E/ in all the eight character locations.

EBLOCK_T<64:0>
: 65 bit vector to be sent to the LDPC encoder containing /E/ in all the eight character locations.

IFRAME_R<71:0>
: 72 bit vector to be sent to the XGMII interface containing /I/ in all the eight character locations

LBLOCK_R<71:0>
: 72 bit vector to be sent to the XGMII interface containing two Local Fault ordered_sets. The Local Fault ordered_set is defined in 46.3.4.

LBLOCK_T<64:0>
: 65 bit vector to be sent to the LDPC encoder containing two Local Fault ordered sets.

55.3.12.2.2 Variables

config
: The config parameter set by PMA and passed to the PCS via the PMA_CONFIG.indicate primitive. Values: MASTER, SLAVE

lber_test_sh
: Boolean variable that is set true when a new LDPC block is available for testing and false when LBER_TEST_SH state is entered. A new LDPC header is available for testing when the Block Sync process has accumulated enough bits from the PMA to evaluate the header of the next block

block_lock
: Boolean variable that is set true when receiver acquires block delineation

hi_lber
: Boolean variable which is asserted true when the lber_cnt exceeds TBD indicating a bit error ratio >TBD

reset
: Boolean variable that controls the resetting of the PCS. It is true whenever a reset is necessary including when reset is initiated from the MDIO, during power on, and when the MDIO has put the PCS into low-power mode.

rx_coded<64:0>
: Vector containing the input to the 64B/65B decoder. The format for this vector is shown in Figure 55–8. The leftmost bit in the figure is rx_coded<0> and the rightmost bit is rx_coded<64>.

rx_raw<71:0>
: Vector containing two successive XGMII transfers. RXC<0> through RXC<3> for the first transfer are placed in rx_raw<0> through rx_raw<3>, respectively. RXC<0> through RXC<3> for the second transfer are placed in rx_raw<4> through rx_raw<7>, respectively. RXD<0> through RXD<31> for the first transfer are placed in rx_raw<8> through rx_raw<39>, respectively. RXD<0> through RXD<31> for the second transfer are placed in rx_raw<40> through rx_raw<71>, respectively.

sh_valid
: Boolean indication that is set true if received LDPC block has valid parity bits.

signal_ok
: Boolean variable that is set based on the most recently received value of PMA_UNITDATA.indicate(SIGNAL_OK). It is true if the value was OK and false if the value was FAIL.

slip_done
: Boolean variable that is asserted true when the SLIP requested by the Block Lock State Machine has been completed indicating that the next candidate frame sync position can be tested.

test_sh
: Boolean variable that is set true when a new sync header is available for testing and false when TEST_SH state is entered. A new sync header is available for testing when the Block Sync process has accumulated enough bits from the PMA to evaluate the header of the next block

tx_coded<64:0>
: Vector containing the output from the 64B/65B encoder. The format for this vector is shown in Figure 55–8. The leftmost bit in the figure is tx_coded<0> and the rightmost bit is tx_coded<64>.

tx_raw<71:0>
: Vector containing two successive XGMII transfers. TXC<0> through TXC<3> for the first transfer are placed in tx_raw<0> through tx_raw<3>, respectively. TXC<0> through TXC<3> for the second transfer are placed in tx_raw<4> through tx_raw<7>, respectively. TXD<0> through TXD<31> for the first transfer are placed in tx_raw<8> through tx_raw<39>, respectively. TXD<0> through TXD<31> for the second transfer are placed in tx_raw<40> through tx_raw<71>, respectively.

55.3.12.2.3 Functions

DECODE(rx_coded<64:0>)

In the PCS Receive process, this function takes as its argument TBD values of rx_symb_vector from the PMA and decodes the 65B-LDPC bit vector returning TBD vectors rx_raw<71:0> which is sent to the XGMII. The DECODE function shall decode the block as specified in 55.3.4.

ENCODE(tx_raw<71:0>)

Encodes the 72-bit vector received from the XGMII, returning TBD values of tx_symb_vector. The ENCODE function shall encode the block as specified in 55.3.4.

R_BLOCK_TYPE = {C, S, T, D, E}

This function classifies each 65-bit rx_coded vector as belonging to one of the five types depending on its contents.

Values: C; The vector contains a sync header of 1 and one of the following:
- a) A block type field of 0x1e and eight valid control characters other than /E/;
- b) A block type field of 0x2d or 0x4b, a valid O code, and four valid control characters;
- c) A block type field of 0x55 and two valid O codes.

S; The vector contains a sync header of 1 and one of the following:
- a) A block type field of 0x33 and four valid control characters;
- b) A block type field of 0x66 and a valid O code;
- c) A block type field of 0x78.

T; The vector contains a sync header of 1, a block type field of 0x87, 0x99, 0xaa, 0xb4, 0xcc, 0xd2, 0xe1 or 0xff and all control characters are valid.

D; The vector contains a sync header of 0.

E; The vector does not meet the criteria for any other value.

A valid control character is one containing a 10GBASE-R control code specified in Table 55–1. A valid O code is one containing an O code specified in Table 55–1.

R_TYPE(rx_coded<64:0>)

Returns the R_BLOCK_TYPE of the rx_coded<64:0> bit vector.

R_TYPE_NEXT

Prescient end of packet check function. It returns the R_BLOCK_TYPE of the rx_coded vector immediately following the current rx_coded vector.

SLIP

Causes the next candidate 65B-LDPC block sync position to be tested. The precise method for determining the next candidate block sync position is not specified and is implementation dependent. However, an implementation shall ensure that all possible bit positions are evaluated.

T_BLOCK_TYPE = {C, S, T, D, E}

This function classifies each 72-bit tx_raw vector as belonging to one of the five types depending on its contents.

Values: C; The vector contains one of the following:
- a) eight valid control characters other than /O/, /S/, /T/ and /E/;
- b) one valid ordered_set and four valid control characters other than /O/, /S/ and /T/;
- c) two valid ordered sets.

S; The vector contains an /S/ in its first or fifth character, any characters before the S character are valid control characters other than /O/, /S/ and /T/ or form a valid ordered_set, and all characters following the /S/ are data characters.

T; The vector contains a /T/ in one of its characters, all characters before the /T/ are data characters, and all characters following the /T/ are valid control characters other than /O/, /S/ and /T/.
D; The vector contains eight data characters.
E; The vector does not meet the criteria for any other value.

A tx_raw character is a control character if its associated TXC bit is asserted. A valid control character is one containing an XGMII control code specified in Table 55–1. A valid ordered_set consists of a valid /O/ character in the first or fifth characters and data characters in the three characters following the /O/. A valid /O/ is any character with a value for O code in Table 55–1.

T_TYPE(tx_raw<71:0>)
Returns the T_BLOCK_TYPE of the tx_raw<71:0> bit vector.

T_TYPE_NEXT
Prescient end of packet check function. It returns the FRAME_TYPE of the tx_raw vector immediately following the current tx_raw vector.

55.3.12.2.4 Counters

lber_cnt
Count up to a maximum of TBD of the number of invalid LDPC codeword within the current TBD µs period.

sh_cnt
Count of the number of sync headers checked within the current TBD LDPC frame window.

sh_invalid_cnt
Count of the number of invalid sync headers within the current TBD LDPC frame window.

55.3.12.2.5 Timers

symb_timer
Continuous timer: The condition symb_timer_done becomes true upon timer expiration.
Restart time: Immediately after expiration; timer restart resets the condition symb_timer_done.
Duration: TBD ns nominal. (See clock tolerance in 55.5.6.2.)

Symb-timer shall be generated synchronously with TX_TCLK. In the PCS Transmit state diagram, the message PMA_UNITDATA.request is issued concurrently with symb_timer_done.

55.3.12.3 Messages

PMA_UNITDATA.indicate (rx_symb_vector)
A signal sent by PMA Receive indicating that a vector of four symbols is available in rx_symb_vector. (See 55.2.6.)

PMA_UNITDATA.request (tx_symb_vector)
A signal sent to PMA Transmit indicating that a vector of four M-ary symbols is available in tx_symb_vector. (See 55.2.5.)

PUDI
Alias for PMA_UNITDATA.indicate (rx_symb_vector).

PUDR
Alias for PMA_UNITDATA.request (tx_symb_vector).

STD

Alias for symb_timer_done.

55.3.12.4 State diagrams

The Lock state machine shown in Figure 55–13 determines when the PCS has obtained lock to the received data stream. The LBER Monitor state machine shown in Figure 55–14 monitors the received signal for high LDPC block error ratio.

The Transmit state machine shown in Figure 55–15 controls the encoding of 65B transmitted blocks. It makes exactly one transition for each 65B transmit block processed. Though the Transmit state machine sends Local Fault ordered sets when reset is asserted, the scrambler and 65B-LDPC may not be operational during reset. Thus, the Local Fault ordered sets may not appear on the PMA service interface.

The Receive state machine shown in Figure 55–16 controls the decoding of received blocks. It makes exactly one transition for each receive block processed.

The PCS shall perform the functions of Lock, LBER Monitor, Transmit and Receive as specified in these state machines.

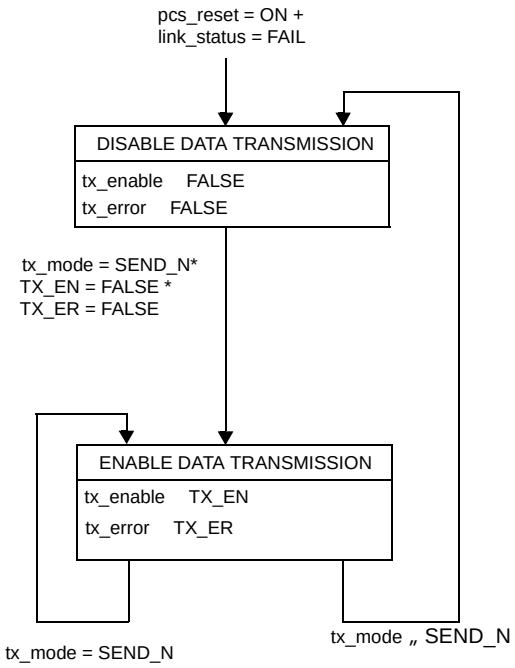


Figure 55–12—PCS Data Transmission Enabling state diagram

55.3.13 PCS Management

The following objects apply to PCS management. If an MDIO Interface is provided (see Clause 45), they are accessed via that interface. If not, it is recommended that an equivalent access be provided.

55.3.13.1 Status

PCS_status:

Indicates whether the PCS is in a fully operational state. It is only true if block_lock is true and

hi_lber is false. This status is reflected in MDIO register 3.32.12. A latch low view of this status is reflected in MDIO register 3.1.2 and a latch high of the inverse of this status, Receive fault, is reflected in MDIO register 3.8.10.

block_lock:

Indicates the state of the block_lock variable. This status is reflected in MDIO register 3.32.0. A latch low view of this status is reflected in MDIO register 3.33.15.

hi_lber:

Indicates the state of the hi_lber variable. This status is reflected in MDIO register 3.32.1. A latch high view of this status is reflected in MDIO register 3.33.14.

55.3.13.2 Counters

The following counters are reset to zero upon read and upon reset of the PCS. When they reach all ones, they stop counting. Their purpose is to help monitor the quality of the link.

lber_count:

TBD-bit counter that counts each time LBER_BAD_SH state is entered. This counter is reflected in MDIO register bits 3.33.13:8. Note that this counter counts a maximum of TBD counts per TBD μs since the LBER_BAD_SH can be entered a maximum of TBD times per TBD μs window.

errored_block_count:

TBD-bit counter. When the receiver is in normal mode, errored_block_count counts once for each time RX_E state is entered. This counter is reflected in MDIO register bits 3.33.7:0.

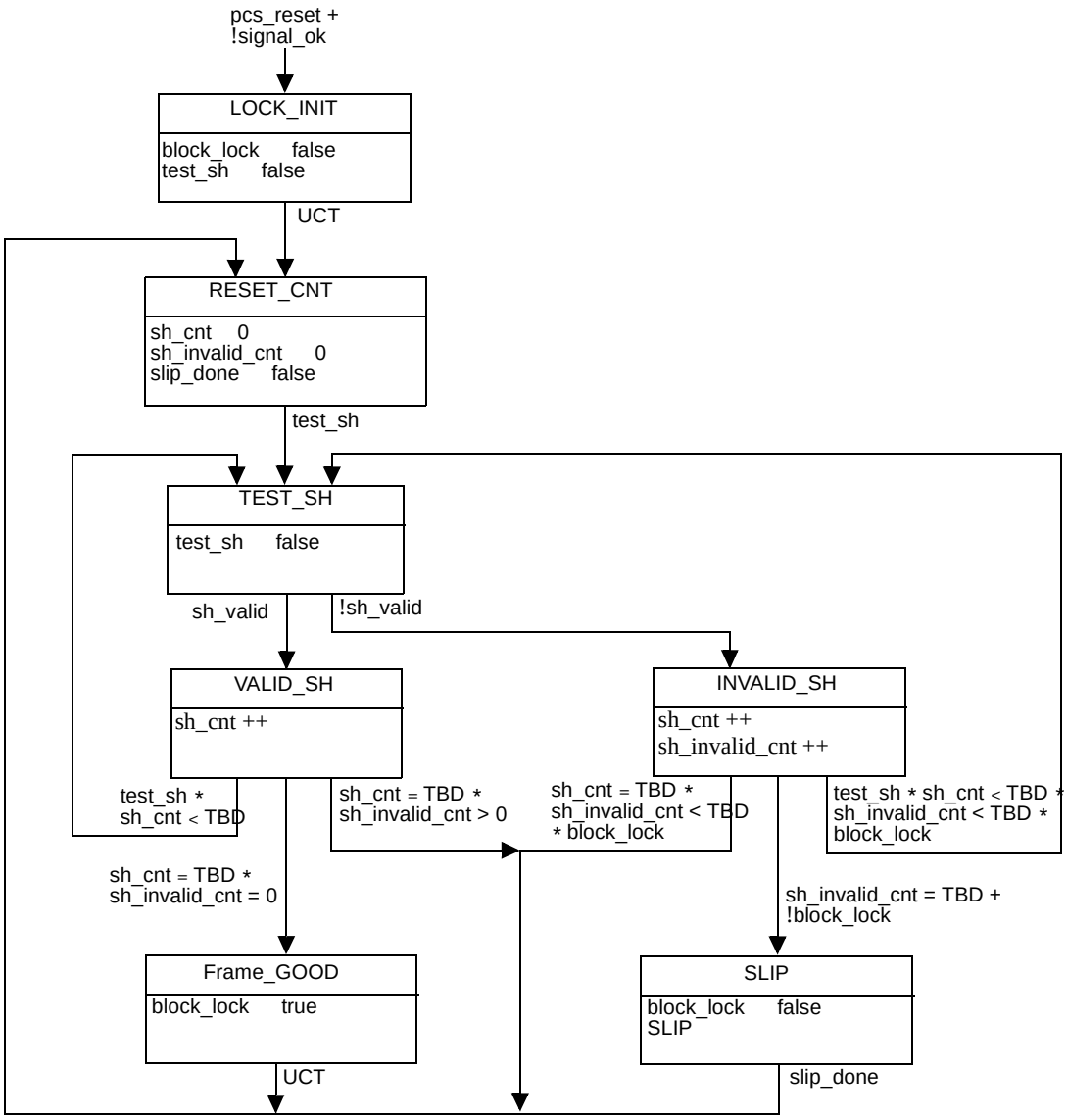


Figure 55–13—Lock state machine

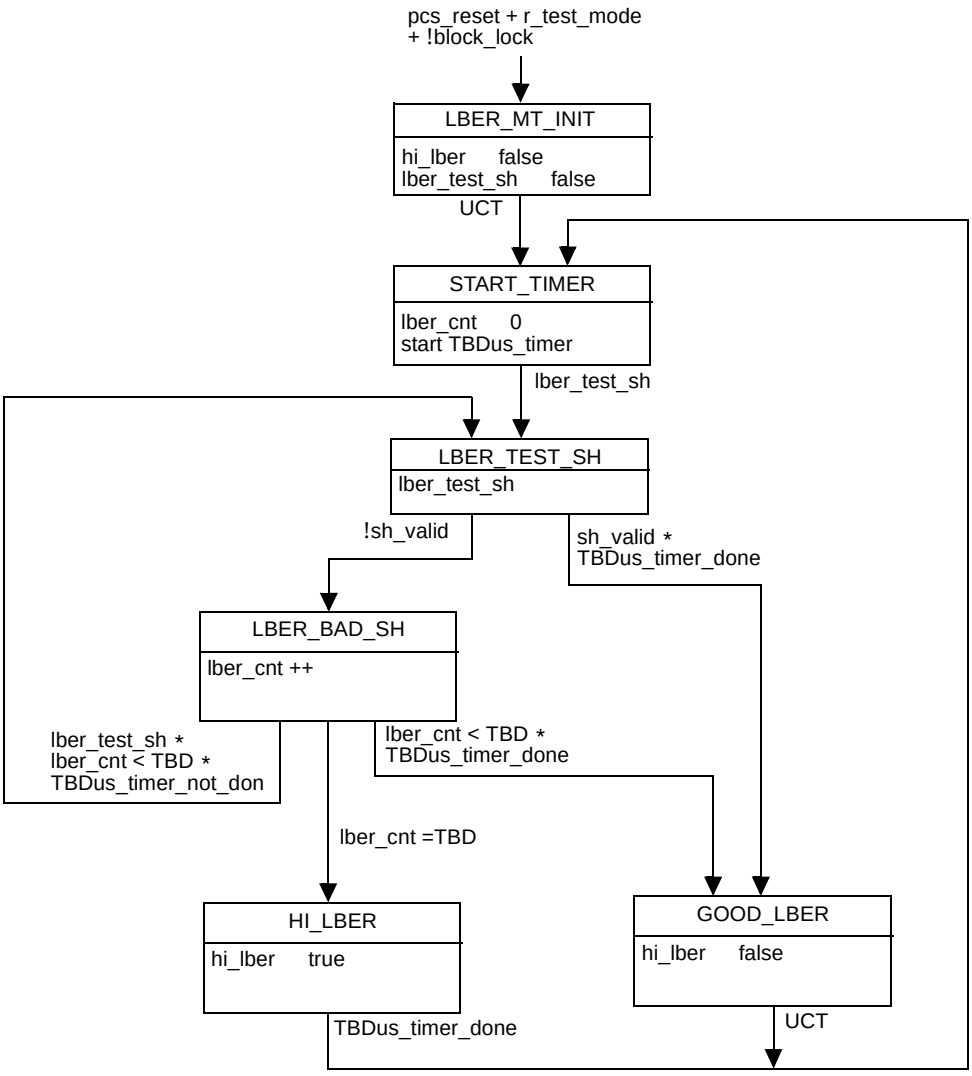


Figure 55–14—LBER monitor state machine

55.3.13.3 Loopback

Editor's note: PCS for 1000BASE-T does not appear to include loop back modes. This can be useful. Please comment.

The PCS shall be placed in Loopback mode when the Loopback bit in MDIO register TBD is set to a logic one. In this mode, the PCS shall accept data on the transmit path from the XGMII and return it on the receive path to the XGMII. In addition, the PCS shall transmit a continuous stream of 0x00FF data words to the PMA sublayer, and shall ignore all data presented to it by the PMA sublayer.

55.4 Physical Medium Attachment (PMA) sublayer

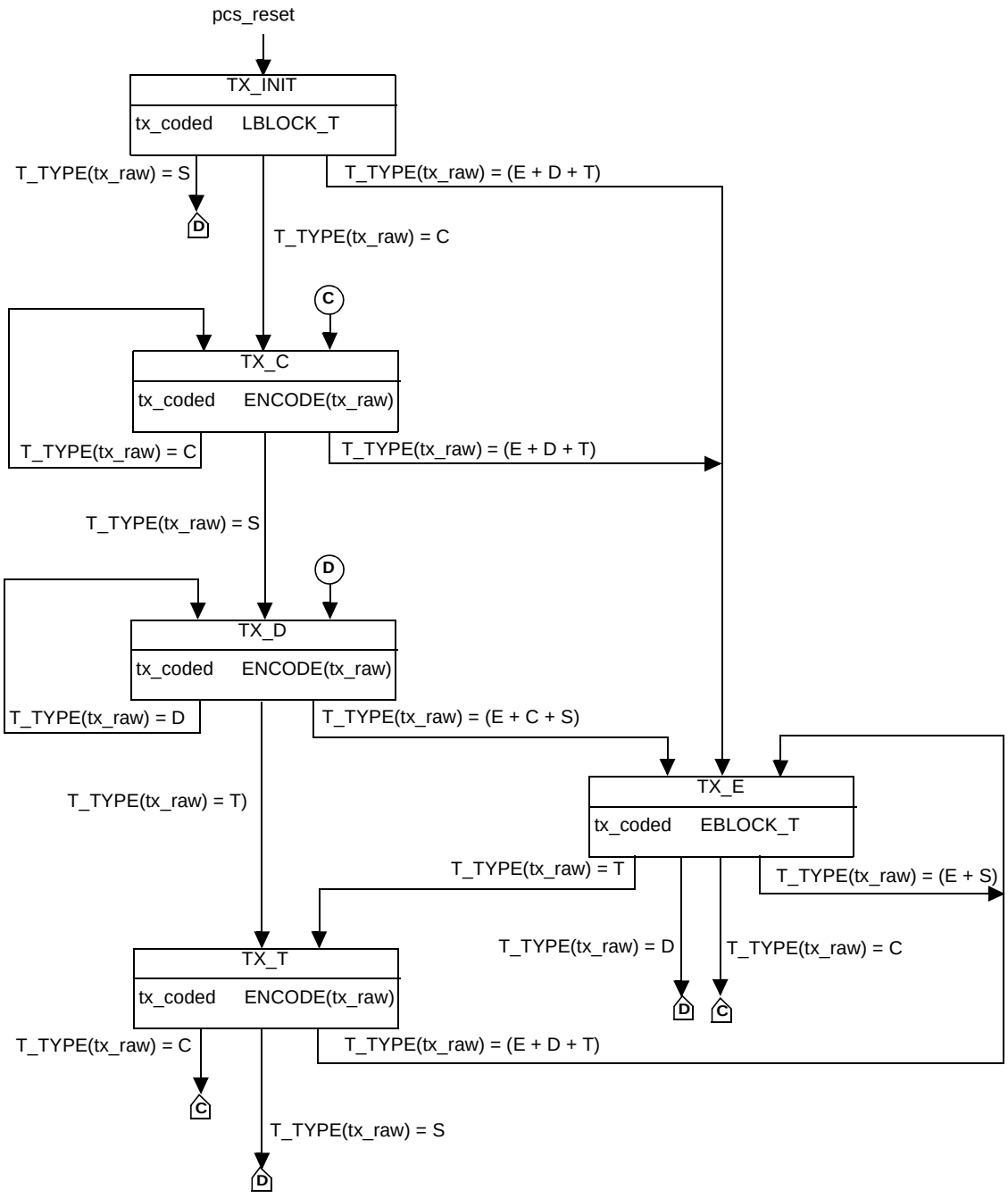


Figure 55–15—PCS Transmit state machine

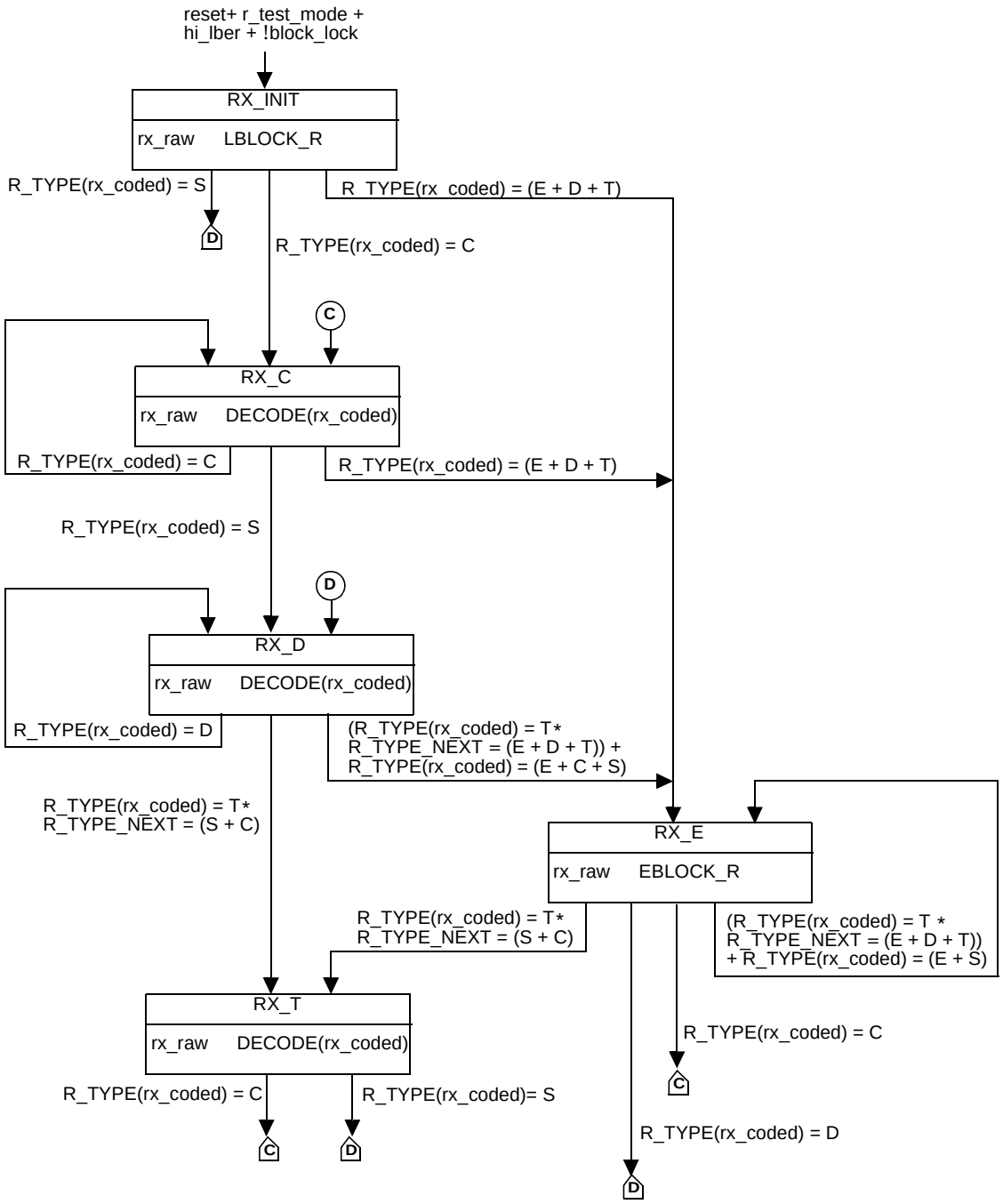


Figure 55–16—PCS Receive state machine

Editor's note: Clause 40 has been used as the starting point for most of the PMA for this initial draft of 10GBASE-T. Please provide comments.

55.4.1 PMA functional specifications

The PMA couples messages from a PMA service interface specified in 55.2.2 to the 10GBASE-T baseband medium, specified in 55.7.

The interface between PMA and the baseband medium is the Medium Dependent Interface (MDI), which is specified in 55.8.

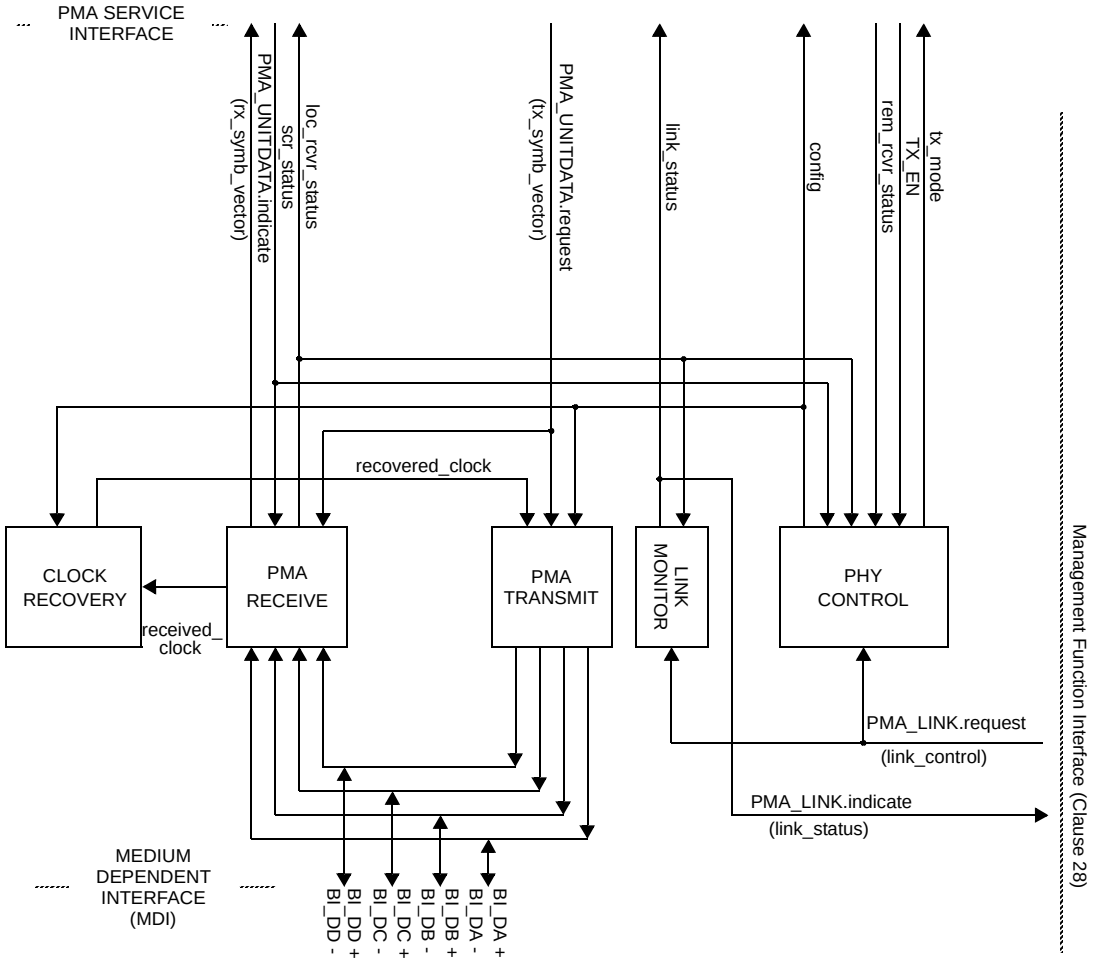


Figure 55–17—PMA reference diagram

NOTE—The recovered_clock arc is shown to indicate delivery of the recovered clock signal back to PMA TRANSMIT for loop timing.

55.4.2 PMA functions

The PMA sublayer comprises one PMA Reset function and five simultaneous and asynchronous operating functions. The PMA operating functions are PHY Control, PMA Transmit, PMA Receive, Link Monitor,

and Clock Recovery. All operating functions are started immediately after the successful completion of the PMA Reset function.

The PMA reference diagram, Figure 55–17, shows how the operating functions relate to the messages of the PMA Service interface and the signals of the MDI. Connections from the management interface, comprising the signals MDC and MDIO, to other layers are pervasive and are not shown in Figure 55–17. The management interface and its functions are specified in Clause 22.

Editor's note: Should we include references to Clause 28 and Clause 30 in addition to Clause 22?

55.4.2.1 PMA Reset function

The PMA Reset function shall be executed whenever one of the two following conditions occur:

a) Power on (see Clause 46.TBD)
b) The receipt of a request for reset from the management entity

PMA Reset sets pcs_reset=ON while any of the above reset conditions hold true. All state diagrams take the open-ended pma_reset branch upon execution of PMA Reset. The reference diagrams do not explicitly show the PMA Reset function.

55.4.2.2 PMA Transmit function

The PMA Transmit function comprises four synchronous transmitters to generate four pulse-amplitude modulated signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD. PMA Transmit shall continuously transmit onto the MDI pulses modulated by the symbols given by tx_symb_vector[BI_DA], tx_symb_vector[BI_DB], tx_symb_vector[BI_DC] and tx_symb_vector[BI_DD], respectively after processing with the THP, TBD transmit filtering, digital to analog conversion (DAC) and subsequent analog filtering. The four transmitters shall be driven by the same transmit clock, TX_TCLK. The signals generated by PMA Transmit shall follow the mathematical description given in 55.4.3.1, and shall comply with the electrical specifications given in 55.5.

When the PMA_CONFIG.indicate parameter config is MASTER, the PMA Transmit function shall source TX_TCLK from a local clock source while meeting the transmit jitter requirements of 55.5.4.1. When the PMA_CONFIG.indicate parameter config is SLAVE, the PMA Transmit function shall source TX_TCLK from the recovered clock of 55.4.2.6 while meeting the jitter requirements of 55.5.4.1.

55.4.2.3 PMA Receive function

The PMA Receive function comprises four independent receivers for pulse-amplitude modulated signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD. PMA Receive contains the circuits necessary to both detect symbol sequences from the signals received at the MDI over receive pairs BI_DA, BI_DB, BI_DC, and BI_DD and to present these sequences to the PCS Receive function. The signals received at the MDI are described mathematically in 55.4.3.2. The PMA shall translate the signals received on pairs BI_DA, BI_DB, BI_DC, and BI_DB into the PMA_UNITDATA.indicate parameter rx_symb_vector with equivalent LBER of less than 10^{-TBD} over a channel meeting the requirements of 55.7.

Editor's note: Following a related concept to 1000BASE-T we have assumed a frame/block error rate instead of a bit error rate measurement. As a reference, Clause 40 specifies FER=1e-7 for packets of size 128bytes.

To achieve the indicated performance, it is highly recommended that PMA Receive include the functions of signal equalization, echo and crosstalk cancellation. The sequence of code-groups assigned to tx_symb_vector is needed to perform echo and self near-end crosstalk cancellation.

The PMA Receive function uses the scr_status parameter and the state of the equalization, cancellation, and estimation functions to determine the quality of the receiver performance, and generates the loc_rcvr_status variable accordingly. The precise algorithm for generation of loc_rcvr_status is implementation dependent.

55.4.2.4 PHY Control function

PHY Control generates the control actions that are needed to bring the PHY into a mode of operation during which frames can be exchanged with the link partner. PHY Control shall comply with the state diagram description given in Figure 55–19 (TBD).

During Auto-Negotiation PHY Control is in the DISABLE 10GBASE-T TRANSMITTER state and the transmitters are disabled. When the Auto-Negotiation process asserts link_control=ENABLE, PHY Control enters the SLAVE SILENT state. Upon entering this state, the maxwait timer is started and PHY Control forces transmission of zeros by setting tx_mode=SEND_Z. The transition out of the SLAVE SILENT state depends on whether the PHY is operating in MASTER or SLAVE mode. In MASTER mode, PHY Control transitions immediately to the PMA TRAINING state. In SLAVE mode, PHY Control transitions to the TRAINING state only after the SLAVE PHY converges its decision feedback equalizer (DFE), acquires timing, and acquires its descrambler state, and sets scr_status=OK**.**

For the SLAVE PHY, the final convergence of the adaptive filter parameters is completed in the TRAINING state. The MASTER PHY performs all its receiver convergence functions in the TRAINING state. Upon entering the PMA TRAINING state, the minwait_timer is started and PHY Control forces transmission into the training mode by asserting tx_mode=SEND_T. After the PHY completes successful training and establishes proper receiver operations, PCS Transmit conveys this information to the link partner via transmission of the parameter loc_rcvr_status (TBD). The link partner's value for loc_rcvr_status is stored in the local device parameter rem_rcvr status. When the minwait_timer expires and the condition loc_rcvr_status=OK is satisfied, PHY Control transitions into TBD.

The normal mode of operation corresponds to the SEND PCS DATA state, where PHY Control asserts tx_mode=SEND_N and transmission of data over the link can take place.

PHY Control may force the transmit scrambler state to be initialized to an arbitrary value by requesting the execution of the PCS Reset function defined in 55.3.2.1.

55.4.2.5 Link Monitor function

Link Monitor determines the status of the underlying receive channel and communicates it via the variable link_status. Failure of the underlying receive channel typically causes the PMA's clients to suspend normal operation.

The Link Monitor function shall comply with the state diagram of Figure 55–20.

Upon power on, reset, or release from power down, the Auto-Negotiation algorithm sets link_control=SCAN_FOR_CARRIER and, during this period, sends fast link pulses to signal its presence to a remote station. If the presence of a remote station is sensed through reception of fast link pulses, the Auto-Negotiation algorithm sets link_control=DISABLE and exchanges Auto-Negotiation information with the remote station. During this period, link_status=FAIL is asserted. If the presence of a remote 10GBASE-T station is established, the Auto-Negotiation algorithm permits full operation by setting link_control=ENABLE. As soon as reliable transmission is achieved, the variable link_status=OK is asserted, upon which further PHY operations can take place.

55.4.2.6 Clock Recovery function

The Clock Recovery function couples to all four receive pairs. It may provide independent clock phases for sampling the signals on each of the four pairs.

The Clock Recovery function shall provide clocks suitable for signal sampling on each line so that the LDPC block-error rate indicated in 55.4.2.3 is achieved. The received clock signal must be stable and ready for use when training has been completed (loc_rcvr_status=OK). The received clock signal is supplied to the PMA Transmit function by received_clock.

55.4.3 MDI

Communication through the MDI is summarized in 55.4.3.1 and 55.4.3.2.

55.4.3.1 MDI signals transmitted by the PHY

The symbols to be transmitted by the PMA on the four pairs BI_DA, BI_DB, BI_DC, and BI_DD are denoted by tx_symb_vector[BI_DA], tx_symb_vector[BI_DB], tx_symb_vector[BI_DC], and tx_symb_vector[BI_DD], respectively. The modulation scheme used over each pair is Pulse Amplitude Modulation. PMA Transmit generates a pulse-amplitude modulated signal on each pair in the following form:

$$b_i = \left(a_i + \sum_k b_{i-k} \mathtt{a}_k \right) mod M$$

$$s(t) = \sum_k b_k h_1(t - kT)$$

In the above equation, a_i represents the M-aryPAM symbol from the set {-(M-1),...,+(M-1)} to be transmitted at time iT, $\mathtt{a}_k$ denotes the THP coefficients, and $h_1(t)$ denotes the system symbol response at the MDI. This symbol response shall comply with the electrical specifications given in 55.5.

55.4.3.2 Signals received at the MDI

Signals received at the MDI can be expressed for each pair as pulse-amplitude modulated signals that are corrupted by noise as follows:

$$r(t) = \sum_k a_{k,agmt} h_2(t - kT) + w(t)$$

In this equation, $h_2(t)$ denotes the impulse response of the overall channel between the transmit symbol source and the receive MDI and $w(t)$ is a term that represents the contribution of various noise sources. The four signals received on pairs BI_DA, BI_DB, BI_DC, and BI_DD shall be processed within the PMA Receive function to yield the received symbols rx_symb_vector.

55.4.4 Automatic MDI/MDI-X Configuration

Automatic MDI/MDI-X Configuration is intended to eliminate the need for crossover cables between similar devices. Implementation of an automatic MDI/MDI-X configuration is required for 10GBASE-T devices. The automatic configuration method used shall comply with the following specifications. The assignment of pin-outs for a 10GBASE-T crossover function cable is shown in Table 55–9 in 55.8.

55.4.4.1 Description of Automatic MDI/MDI-X state machine

The Automatic MDI/MDI-X state machine facilitates switching the BI_DA(C)+ and BI_DA(C)– with the BI_DB(D)+ and BI_DB(D)– signals respectively prior to the auto-negotiation mode of operation so that fast link pulses can be transmitted and received in compliance with Clause 28 Auto-Negotiation specifications. The correct polarization of the crossover circuit is determined by an algorithm that controls the switching function. This algorithm uses an 11-bit Linear Feedback Shift Register (LFSR) to create a pseudo-random sequence that each end of the link uses to determine its proposed configuration. Upon making the selection to either MDI or MDI-X, the node waits for a specified amount of time while evaluating its receive channel to determine whether the other end of the link is sending link pulses or PHY-dependent data. If link pulses or PHY-dependent data are detected, it remains in that configuration. If link pulses or PHY-dependent data are not detected, it increments its LFSR and makes a decision to switch based on the value of the next bit. The state machine does not move from one state to another while link pulses are being transmitted.

55.4.4.2 Pseudo-random sequence generator

One possible implementation of the pseudo-random sequence generator using a linear-feedback shift register is shown in Figure 55–18. The bits stored in the shift register delay line at time n are denoted by S[10:0]. At each sample period, the shift register is advanced by one bit and one new bit represented by S[0] is generated. Switch control is determined by S[10].

Editor's note: This scrambler has been taken from Clause 40. This results in a third different scrambler, in addition to the one from Clause 49 for scrambling the 65B blocks and the second from Clause 40 for PAM training. Please provide comments.

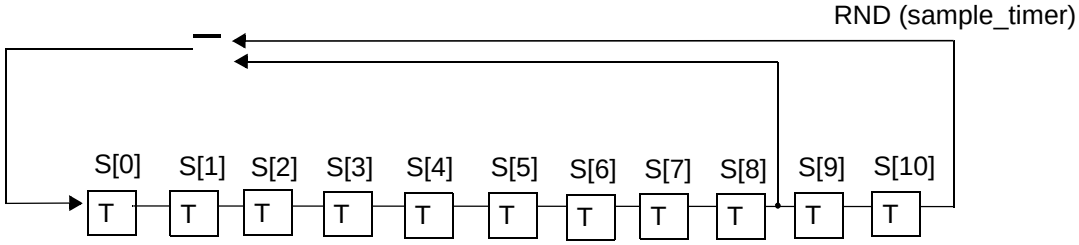


Figure 55–18—Automatic MDI/MDI-X linear-feedback shift register

55.4.5 State variables

55.4.5.1 State diagram variables

config

The PMA shall generate this variable continuously and pass it to the PCS via the PMA_CONFIG.indicate primitive.

Values: MASTER or SLAVE

link_control

This variable is defined in 28.2.6.2.

Link_Det

This variable indicates linkpulse = true or link_status = READY has occurred at the receiver since the last time sample_timer has been started.

Values: TRUE: linkpulse = true or link_status = READY has occurred since the last time sample_timer has been started.

FALSE: otherwise

55.4.5.2 Timers

All timers operate in the manner described in 14.2.3.2 with the following addition. A timer is reset and stops counting upon entering a state where “stop timer” is asserted.

A_timer

An asynchronous (to the Auto-Crossover State Machine) free-running timer that provides for a relatively arbitrary reset of the state machine to its initial state. This timer is used to reduce the probability of a lock-up condition where both nodes have the same identical seed initialization at the same point in time.

Values: The condition A_timer_done becomes true upon timer expiration.

Duration: This timer shall have a period of TBD − TBD%.

Initialization of A_timer is implementation specific.

maxwait_timer

A timer used to limit the amount of time during which a receiver dwells in the SLAVE SILENT and TRAINING states. The timer shall expire TBD − TBD ms if config = MASTER or TBD − TBD ms if config = SLAVE. This timer is used jointly in the PHY Control and Link Monitor stage diagrams. The maxwait_timer is tested by the Link Monitor to force link_status to be set to FAIL if the timer expires and loc_rcvr_status is NOT_OK. See Figure 55–19.

minwait_timer

A timer used to determine the minimum amount of time the PHY Control stays in the PMA TRAINING, SEND IDLE, or DATA states. The timer shall expire TBD − TBDμs after being started.

sample_timer

This timer provides a long enough sampling window to ensure detection of Link Pulses or link_status, if they exist at the receiver.

Values: The condition sample_timer_done becomes true upon timer expiration.

Duration: This timer shall have a period of TBD − TBD ms.

stabilize_timer

A timer used to control the minimum time that loc_rcvr_status must be OK before a transition to Link Up can occur. The timer shall expire TBD − TBD μs after being started.

55.4.6 State Diagrams

55.4.6.1 PHY Control state diagram

Editor’s note: Clause 40 has been used as the starting point for this initial PHY Control state diagram with some additional outline THP states. The THP effect in the PHY start-up requires additional input for the Task Force. Please provide comments.

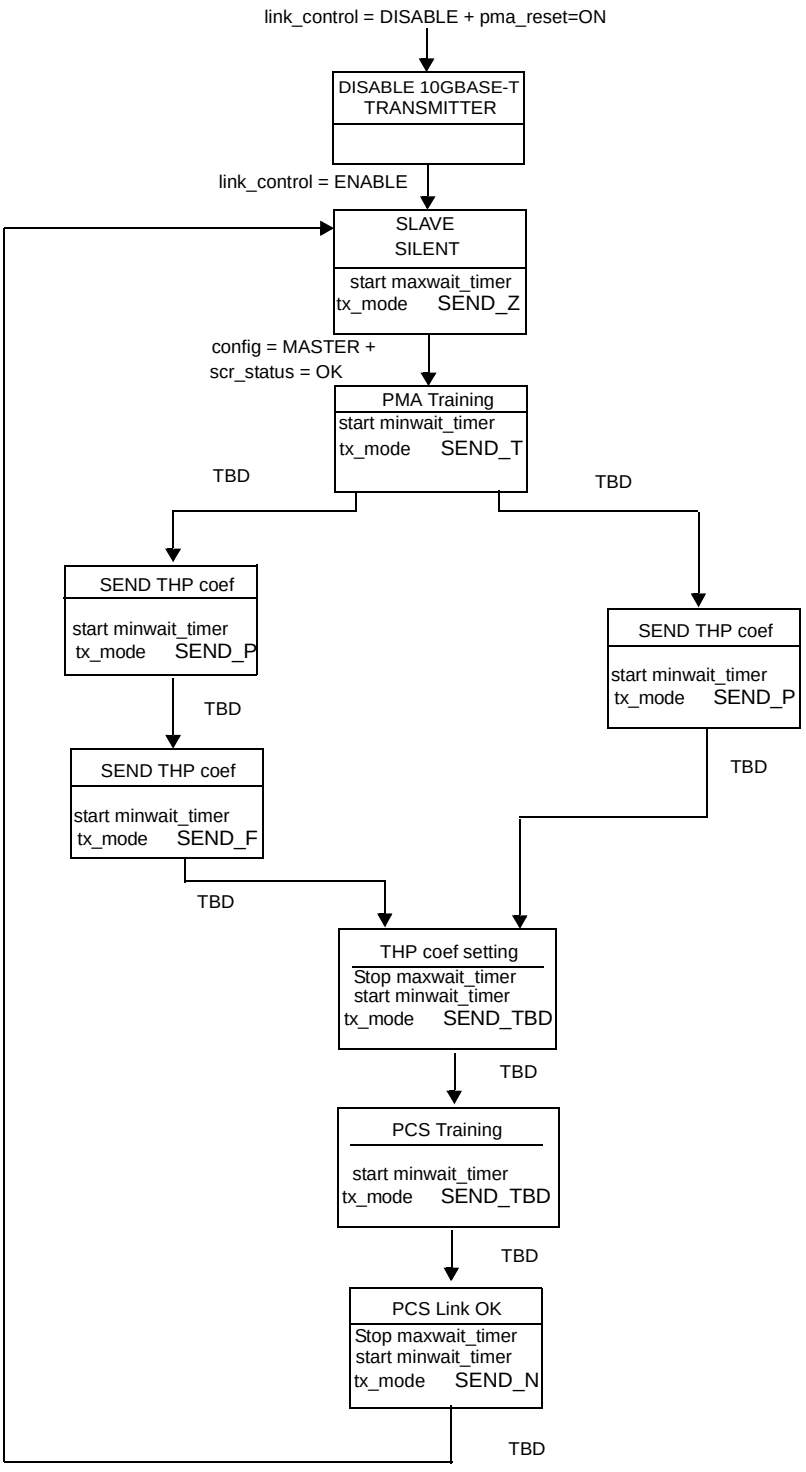
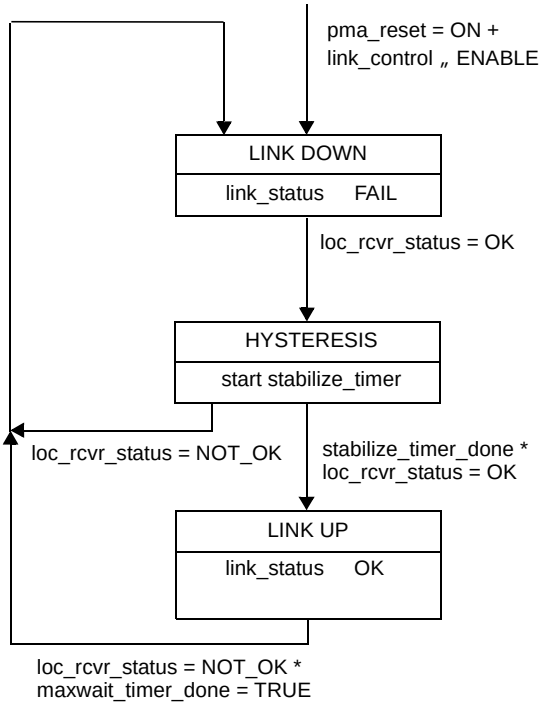


Figure 55–19—TBD PHY Control state diagram

55.4.6.2 Link Monitor state diagram



NOTES
1—maxwait_timer is started in PHY Control state diagram (see Figure 55–19).
2—The variables link_control and link_status are designated as link_control_(10GigT) and link_status_(10GigT), respectively, by the Auto-Negotiation Arbitration state diagram (TBD Figure 28–16).

Figure 55–20—Link Monitor state diagram

55.5 PMA electrical specifications

This subclause defines the electrical characteristics of the PMA and specifies PMA-to-MDI interface tests.

Common-mode tests use the common-mode return point as a reference.

55.5.1 Isolation requirement

The PHY shall provide electrical isolation between the port device circuits, including frame ground (if any) and all MDI leads. This electrical separation shall withstand at least one of the following electrical strength tests:

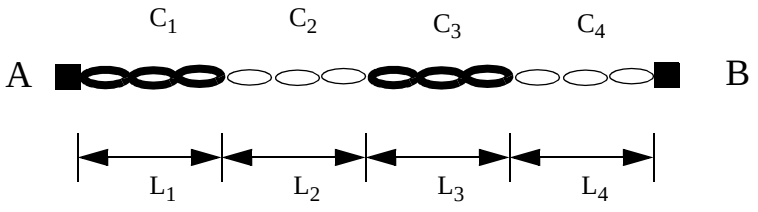
a) 1500 V rms at 50 Hz to 60 Hz for 60 s, applied as specified in Section 5.3.2 of IEC 60950-1: 2001.
b) 2250 Vdc for 60 s, applied as specified in Section 5.3.2 of IEC 60950-1: 2001.
c) A sequence of ten 2400 V impulses of alternating polarity, applied at intervals of not less than 1 s. The shape of the impulses shall be 1.2/50 ms (1.2 ms virtual front time, 50 ms virtual time or half value), as defined in IEC 60060.

There shall be no insulation breakdown, as defined in Section 5.3.2 of IEC 60950-1: 2001, during the test. The resistance after the test shall be at least 2 MW, measured at 500 Vdc.

Editor's note: Text in Clause 55.5.1 copied from 1000BASE-T standard, section 40.6.1.1 with updates to some of the references. Are there additional updates?

55.5.1.1 Test channel

To perform the transmitter MASTER-SLAVE timing jitter tests described in this clause, a test channel is required to ensure that jitter is measured under conditions of poor signal to echo ratio. This test channel shall be constructed by combining 100 W and 120 W cable segments that both meet or exceed ISO/IEC 11801 augmented by link segment specified in Clause 55.7 for each pair, as shown in Figure 55–21, with the lengths and additional restrictions on parameters described in Table 55–2. The ends of the test channel shall be terminated with connectors meeting or exceeding ISO/IEC 11801:2001 Class E specifications. The return loss of the resulting test channel shall meet the return loss requirements of Clause 55.7.2.3 and the crosstalk requirements of Clause 55.7.3.3.



Identical for each of the four pairs.

Figure 55–21—Test channel topology for each cable pair

Table 55–2—Test channel cable segment specifications

Cable segment	Length (meters)	Characteristic impedance (at frequencies > 1 MHz)	Attenuation (per 100 meters at TBD freq)
1	$L_1=x_1$	120 – 5W	TBD
2	$L_2=x$	100 – 5W	TBD
3	$L_3=x_2$	120 – 5W	TBD
4	$L_4=y$	100 – 5W	TBD

Editor's Note: —x_1, x_2, x and y as well as the Attenuation numbers are TBD.

55.5.2 Test modes

Editor's Note: Until the symbol rate is chosen, we use Fs for the symbol rate; Fs is in MHz.

The test modes described below shall be provided to allow for testing of the transmitter waveform, transmitter distortion, and transmitted jitter.

For a PHY with an MDIO management interface, these modes shall be enabled by setting ***TBD*** bits(10GBASE-T Control Register) of the MDIO Management register set as shown in Table-55–3. These test modes shall only change the data symbols provided to the transmitter circuitry and shall not alter the electrical and jitter characteristics of the transmitter and receiver from those of normal (non-test mode) operation. PHYs without a MDIO shall provide a means to enable these modes for conformance testing.

Table 55–3—MDIO management register settings for test modes

Bit 1 TBD	Bit 2 TBD	Bit 3 TBD	Mode
0	0	0	Normal operation
0	0	1	Test mode 1—Transmit waveform test
0	1	0	Test mode 2—Transmit jitter test in MASTER mode
0	1	1	Test mode 3—Transmit jitter test in SLAVE mode
1	0	0	Test mode 4—Transmit distortion test
1	0	1	Reserved, operations not identified.
1	1	0	Reserved, operations not identified.
1	1	1	Reserved, operations not identified.

When test mode 1 is enabled, the PHY shall transmit the following sequence of data symbols A_n, B_n, C_n, D_n, of 55.4.3.1 continually from all four transmitters, with the THP turned off:

{{3 +M followed by 125 0 symbols}, {3 –M followed by 125 0 symbols},{3 +(M/2) followed by 125 0 symbols},{3 –(M/2) followed by 125 0 symbols}, {128 +M symbols, 128 –M symbols, 128 +M symbols, 128 –M symbols}, {4096 0 symbols}}

This sequence is repeated continually without breaks between the repetitions when the test mode is enabled. A typical transmitter output is shown in Figure 55–22. The transmitter shall time the transmitted symbols from a Fs – 0.01% clock in the MASTER timing mode.

Editor's Note: For PAM12, M=12, for PAM8, M=8. The above sequence has been adapted from the sequence used in 1000BASE-T. Reasons for this choice are provided in the editor's note in 55.5.3.1, when peak differential output and level accuracy are discussed. This sequence has not been discussed by the task force and alternate proposals for the sequence are welcome. Power back-off of the transmitter when transmitting data over shorter cable lengths has been discussed in the task force. When an appropriate methodology and numbers are agreed upon in the task force, some/all of the test modes described herein may have to be modified/enhanced by additional test modes to test the transmitter.

When test mode 2 is enabled, the PHY shall transmit, with THP turned off, the data symbol sequence, {+M, –M} repeatedly on all channels. The transmitter shall time the transmitted symbols from an Fs – 0.01% clock in the MASTER timing mode.

When test mode 3 is enabled, the PHY shall transmit the data symbol sequence {+M, -M} repeatedly on all channels. The transmitter shall time the transmitted symbols from an Fs – 0.01% clock in the SLAVE timing mode. A typical transmitter output for transmitter test modes 2 & 3 in is shown in Figure 55–23.

Editor's note: The waveforms shown in Figure 55–22 and Figure 55–23 assume a symbol rate of 800MHz and correspond to a peak-to-peak output swing of 2V differential. These are illustrative waveforms and can be adjusted once both the parameters are finalized by the task force.

When test mode 4 is enabled, the PHY shall transmit, with the THP turned off, transmitted symbols, timed from a Fs – 0.01% clock in the MASTER timing mode, defined as follows:

Symbols corresponding to a single frequency tone, with frequencies of (Fs/1024)*13, (Fs/1024)*23, (Fs/1024)*53, (Fs/1024)*101, (Fs/1024)*167.

Symbols corresponding to dual frequency tones in the pairs of [(Fs/1024)*179, (Fs/1024)*181], [(Fs/1024)*277, (Fs/1024)*281], [(Fs/1024)*397, (Fs/1024)*401], [(Fs/1024)*499, (Fs/1024)*503],

The peak to peak symbols used in this test, for both single and dual frequency tones correspond to – M .

Editor's Note: Please provide comments and/or alternatives on the above specification for test mode 4. As written, this deviates from the 1000BASE-T approach because time domain distortion tests based on a scrambled sequence will be hard/marginal to measure with currently available test equipment. This is because of the higher speed and the lower distortion specifications. This issue was discussed in detail in http://www.ieee802.org/3/an/public/jul04/gupta_1_0704.pdf *at the July 2004 meeting.*

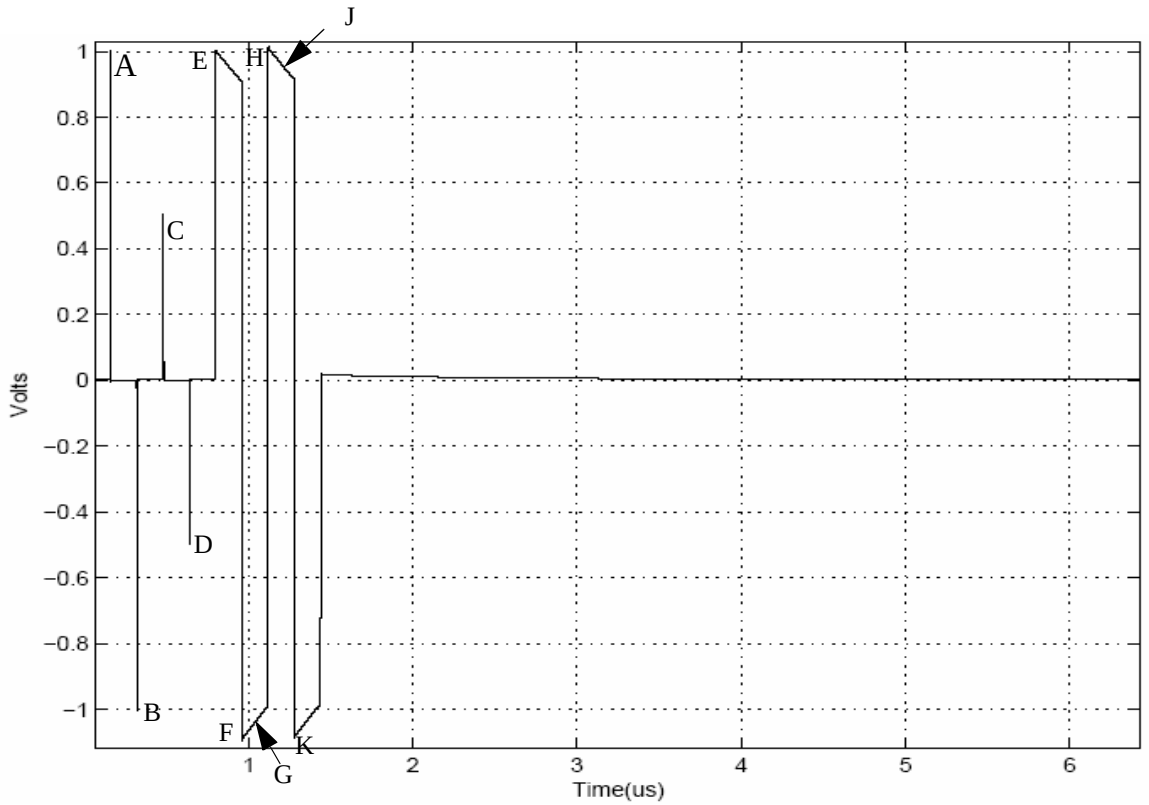


Figure 55–22—Example of transmitter test mode 1 waveform (1 cycle): informative

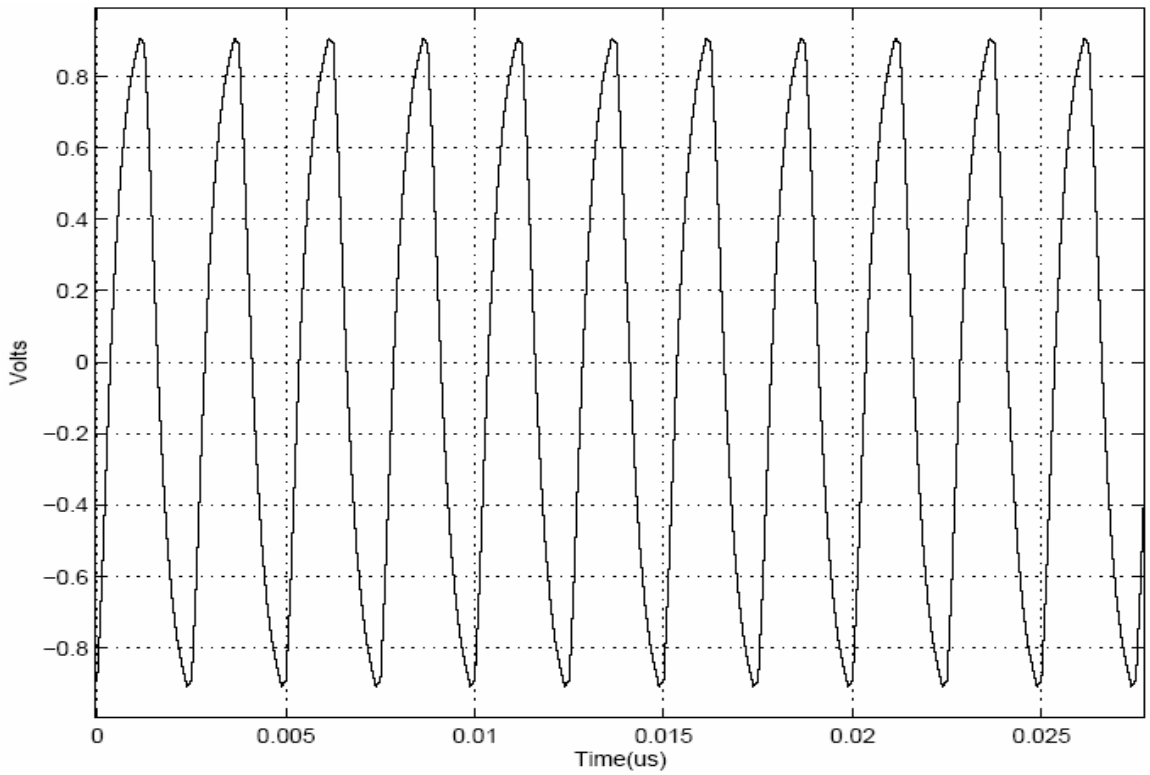


Figure 55–23—Example of transmitter test modes 2 & 3 waveform: informative

55.5.2.1 Test Fixtures

The following fixtures (illustrated by Figure 55–24, Figure 55–25 and Figure 55–26), or their functional equivalents, shall be used for measuring the transmitter specifications described in 55.5.3.

Editor's note: The next three figures have been changed from the 1000BASE-T test setups in places where deemed necessary. Figure 55–25 includes a power summer device to subtract the outputs of the transmitter and provide correct impedance transformation from the 50 W single-ended output impedances to a 50 W Spectrum Analyzer input. Comments/alternate proposals are welcome.

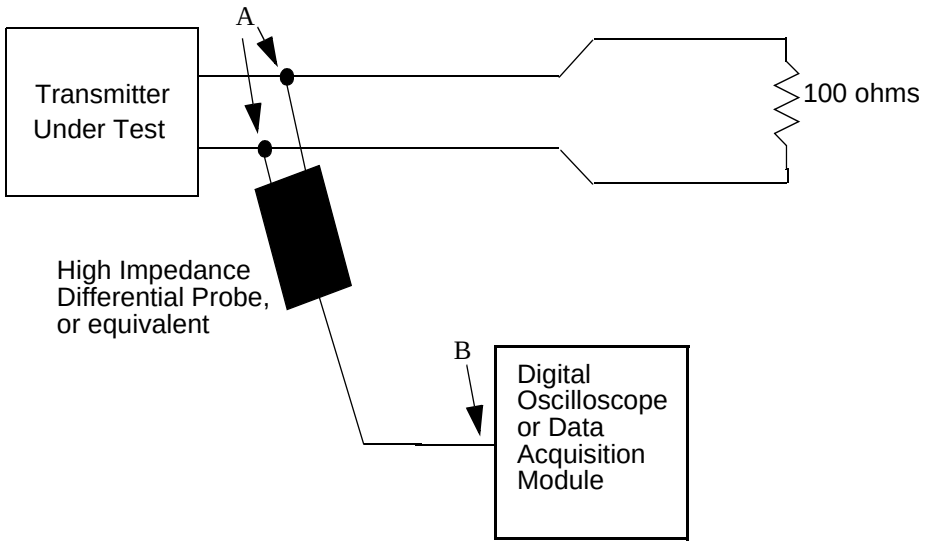


Figure 55–24—Transmitter test fixture 1 for template and droop measurement

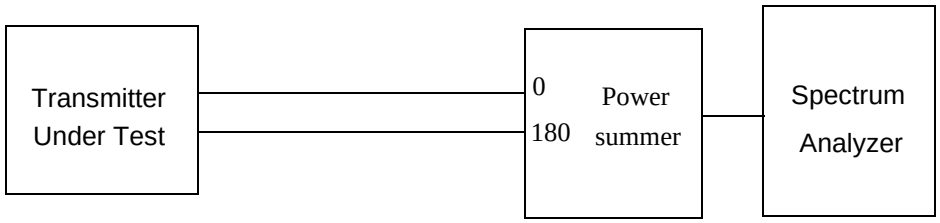


Figure 55–25—Transmitter test fixture 2 for linearity measurement

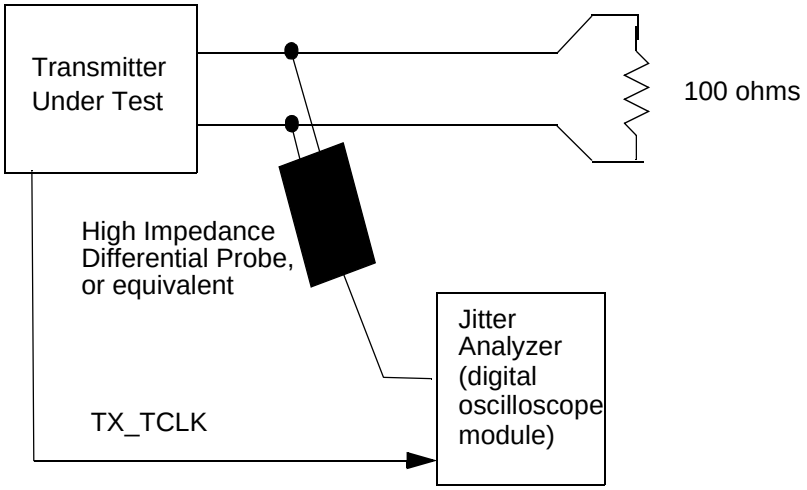


Figure 55–26—Transmitter test fixture 3 for transmitter jitter measurement

To allow for measurement of transmitted jitter in test mode 2 and 3, for both master and slave modes, the PHY shall provide access to the Fs rate symbol clock, TX_TCLK, that times the transmitted symbols (see 55.4.2.2). The PHY shall provide a means to enable this clock output if it is not normally enabled.

55.5.3 Transmitter electrical specifications

The PMA shall provide the Transmit function specified in 55.4.2.2 in accordance with the electrical specifications of this clause.

Where a load is not specified, the transmitter shall meet the requirements of this clause with a 100 W resistive differential load connected to each transmitter output.

55.5.3.1 Peak differential output voltage and level accuracy

The peak value of the waveform at points A, B, C, D, E, F, H indicated in Figure 55–22 is defined by measuring the peak of the absolute value of the waveform between the time interval of [2/Fs 3/Fs] after the start time of the corresponding non-zero symbol ($-M$, $-M/2$)). The time instants corresponding to these peak measured values are defined as t_A, t_B, t_C, t_D, t_E, t_F, t_H

The peak value of the waveform at points A and B, as defined above, shall fall within the range of 1V −5% .

These measurements are to be made for each pair while operating in test mode 1 and observing the differential signal output at the MDI using transmitter test fixture 1 with no intervening cable.

The peak of the waveforms at points A and B shall differ by less than 1%.

The peak of the waveform at points C and D as defined above shall differ by less than 2% from 0.5 times the average of the peaks of the waveform at points A and B.

Editor's note: The peak value is defined in such a way for two reasons. Firstly the waveform may not settle to the correct peak in one baud period with the recommended bandwidth specification for the transmitter. Secondly the magnetics may exhibit a significant insertion loss(IL) variation at high frequencies, and have a well controlled IL variation only at low frequencies.

The range is chosen to be tighter at −5% as opposed to −10% in 1000BASE-T. This is to reduce the possible significant margin loss in the system, which would occur if a transmitter transmits a 10% lower voltage and the receiver receiving the data of such a transmitter, is in presence of ANEXT interferers which are 10% higher. The peak voltage must be between 1V and 1.25V based on the motion in

http://www.ieee802.org/3/an/public/jul04/motions_1_0704.pdf

The peak voltage is chosen at the lower end, i.e. at 1V, as a suggested value in this draft, for various reasons discussed in the following presentations:

http://www.ieee802.org/3/an/public/may04/gupta_1_0504.pdf

http://www.ieee802.org/3/an/public/jul04/gupta_1_0704.pdf

Level matching accuracy requirements are kept similar to 1000BASE-T, the linearity of the transmitter is defined separately as an overriding test on the level matching accuracy. Comments/Suggestions on the numbers/methodology discussed in this clause are welcome.

55.5.3.2 Maximum output droop

The magnitude of the value of the waveform at point G, as defined in Figure 55–22, shall be greater than (100-5.5*800/Fs)% of the peak value of the waveform at point F as defined in 55.5.3.1. These measurements are to be made for each pair while in test mode 1 and observing the differential signal output at the MDI using transmit test fixture 1 with no intervening cable. Point G is defined as the waveform at exactly (64/Fs) time after time instant t_F. Additionally, the magnitude of the value of the waveform at point J as defined in Figure 55–22 shall be greater than (100-5.5*800/Fs)% of the magnitude of the peak value of the waveform at point H, as defined in 55.5.3.1. Point J is defined as the waveform at exactly (64/Fs) time after time instant t_H.

Editor's note: The numbers described above are similar to those discussed in the presentation:

http://www.ieee802.org/3/an/public/jul04/gupta_1_0704.pdf

The droop is chosen with 10% margin for a 100kHz transformer high pass pole. Alternatives proposals are welcome.

55.5.3.3 Differential output templates

The voltage waveforms around points A, B, C, D defined in Figure 55–22, after the normalization described herein, shall lie within the time domain template 1 defined in Figure 55–27 and the piecewise linear interpolation between the points in Table-55–4. These measurements are to be made for each pair while in test mode 1 and while observing the differential signal output at the MDI using transmitter test fixture 1 with no intervening cable. The waveforms may be shifted in time as appropriate to fit within the template.

The waveform around point A is normalized by dividing by the peak value of the waveform at A.

The waveform around point B is normalized by dividing by the negative of the peak value of the waveform at A.

The waveform around point C is normalized by dividing by 1/2 the peak value of the waveform at A.

The waveform around point D is normalized by dividing by the negative of 1/2 the peak value of the waveform at A.

The voltage waveforms around points F and H defined in Figure 55–22, after the normalization described herein, shall lie within the time domain template 2 defined in Figure 55–27 and the piecewise linear interpolation between the points in Table-55–5. These measurements are to be made for each pair while in test mode 1 and while observing the differential signal output at the MDI using transmitter test fixture 1 with no intervening cable. The waveforms may be shifted in time as appropriate to fit within the template.

The waveform around point F is normalized by dividing by the peak value of the waveform at F.

The waveform around point H is normalized by dividing by the peak value of the waveform at H.

*Editor's Note: The templates are to be created with the following assumptions about the elements in the transmit path. Deviation from 1000BASE-T include the elimination of the digital filter, and the elimination of the 2MHz high pass test filter. The -3dB frequency of the two pole continuous time filter is chosen to be 0.9*Fs/2 to 0.75*Fs mainly because the transformer will not have significant excess bandwidth because of the presence of an additional pole at the output. Comments are welcome.*

1. Ideal DAC.

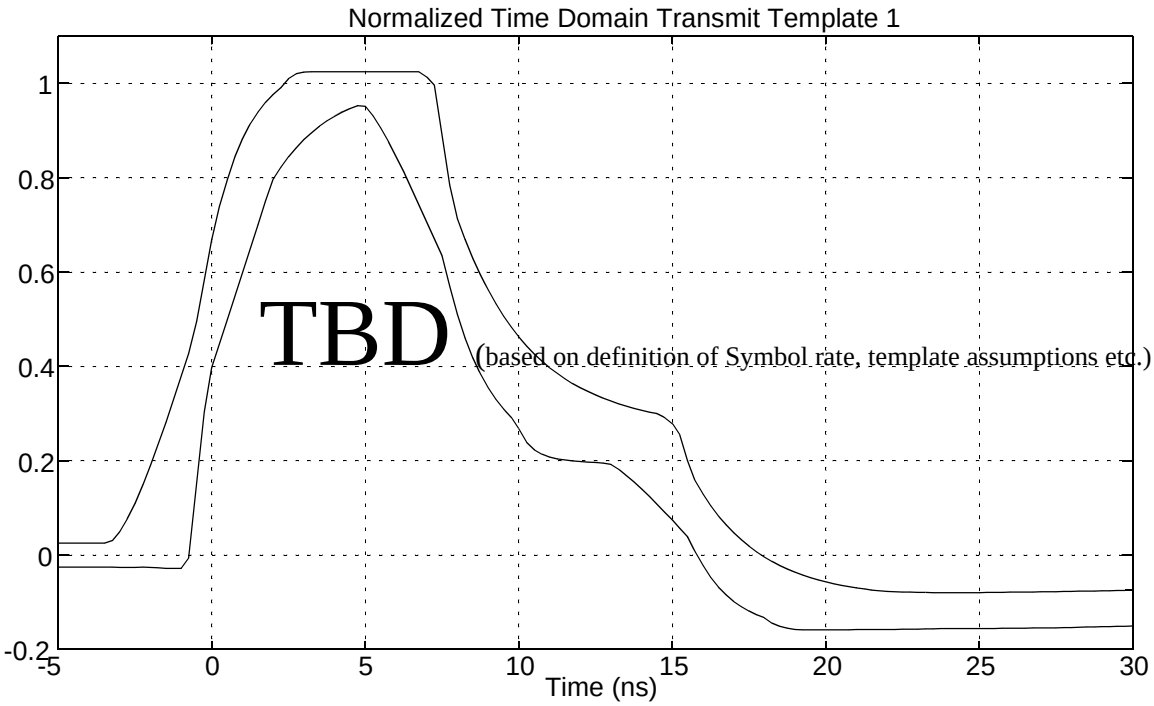
2. Two pole continuous time low pass filter with -3dB frequency varying from 0.9*Fs/2 to 0.75Fs.

3. Single pole continuous time high pass filter (transformer high pass) with pole varying from 1Hz to 100kHz.

4. Additionally, +0.025 to be added to the upper template and –0.025 to be added to the lower template to allow for noise and measurement error.

Editor's note: The Figure 55–27 and Table-55–4 and Table-55–5 have been borrowed from 1000BASE-T and are place holders in this draft. The templates/tables can be generated when the above assumptions are accepted by the Task Force - either as is, or with modifications.

NOTE—The transmit templates are not intended to address electromagnetic radiation limits.



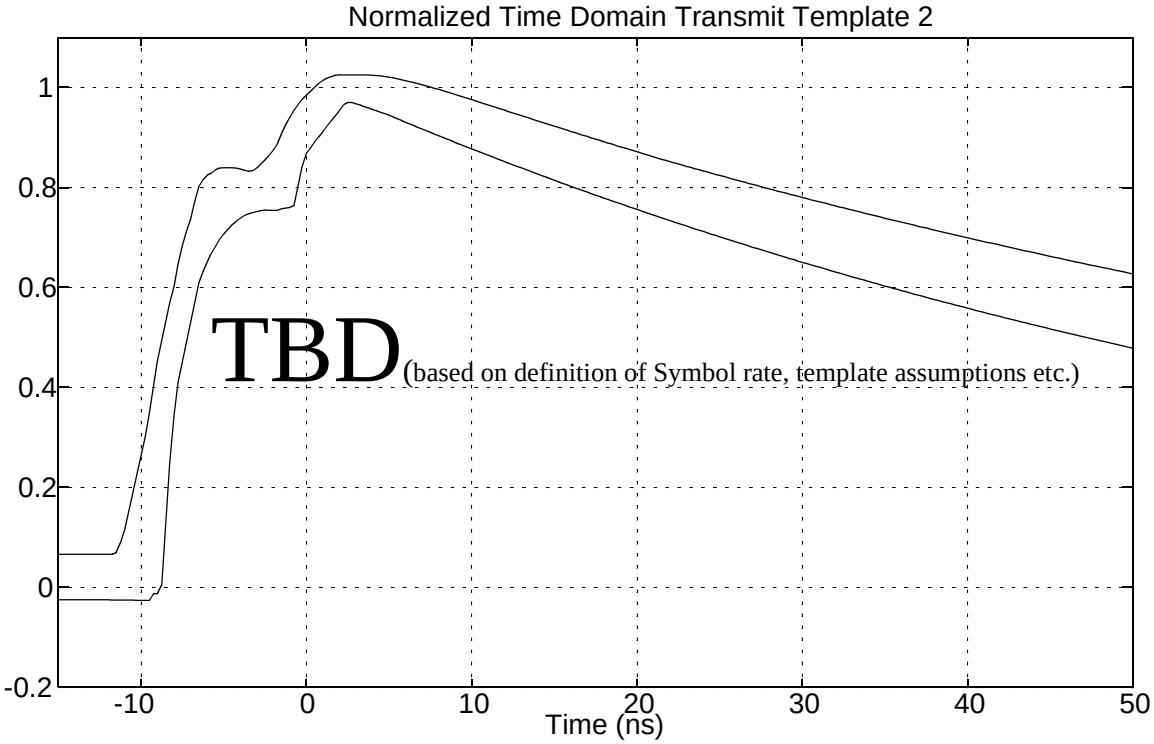


Figure 55–27—Normalized transmit templates as measured at MDI using transmit test fixture 1

Table 55–4—Normalized time domain voltage template 1

Time, ns	Normalized transmit time domain template, upper limit	Normalized transmit time domain template, lower limit	Time, ns	Normalized transmit time domain template, upper limit	Normalized transmit time domain template, lower limit
–5.00	tbd	–tbd	tbd	tbd	tbd

Table 55–5—Normalized time domain voltage template 2

Time, ns	Normalized transmit time domain template, upper limit	Normalized transmit time domain template, lower limit	Time, ns	Normalized transmit time domain template, upper limit	Normalized transmit time domain template, lower limit
–15.00	tbd	tbd	tbd	tbd	tbd

55.5.4 Transmitter linearity.

When in test mode 4 and observing the spectrum of the differential signal output at the MDI using transmitter test fixture 2, for each pair, with no intervening cable, the transmitter linearity mask to be defined as follows.

The SFDR of the transmitter when subject to single tone inputs producing output with peak to peak transmit amplitude shall be:

better than X_{nonlin} dB in the frequency range, $f \in (0.1, f_1]$ MHz, f_1 is in MHz

and better than $[X_{nonlin}-X_{nlslope}*\log10(f/f_1)]$ dB, for $f \in (f_1, Fs/6]$ MHz.

The Signal to Intermodulation distortion ratio of the transmitter, for dual tone inputs, producing output with peak to peak transmit amplitude, shall be better than:

$[X_{nonlin} +2.5- X_{nlslope}*\log10(f/f_1)]$ dB for $f \in (Fs/6, Fs/2]$MHz

The specification on transmit linearity, is provided for the interoperability of the far end device. As a normative specification, the parameter X_{nonlin} = 65dB, parameter f_1= 25MHz, and parameter $X_{nlslope}$ =20dB. The recommended specification is X_{nonlin} = 68dB and parameter $X_{nlslope}$ =0dB.

Editor's note: The above specification assumes a frequency domain approach to measuring transmitter non-linearity. The rationale for taking the frequency domain approach was explained in section 55.5.2. Alternate proposals to this methodology/numbers chosen above are welcome.

The "normative" and "recommended" specification is provided in accordance with the motion in the IEEE July meeting.

http://www.ieee802.org/3/an/public/jul04/motions_1_0704.pdf

The recommended specification on transmitter linearity will enable the local receiver to achieve the echo cancellation required to meet its BER performance requirements without need for non-linear cancellers and external hybrids. If the transceiver uses other techniques to suppress the impact of local transmitter nonlinearity and can meet its performance requirements in their presence, compliance with the recommended linearity requirement is not required.

55.5.4.1 Transmitter timing jitter

When in test mode 2 or test mode 3, the peak-to-peak jitter J_{txout} of the zero crossings of the differential signal output at the MDI relative to the corresponding edge of TX_TCLK is measured. The corresponding edge of TX_TCLK is the edge of the transmit test clock, in polarity and time, that generates the zero-crossing transition being measured.

When in the normal mode of operation as the MASTER, the peak-to-peak value of the MASTER TX_TCLK jitter relative to an unjittered reference shall be less than J_1. When the jitter waveform on TX_TCLK is filtered by a high-pass filter, $H_{jf1}(f)$, having the transfer function below, the peak-to-peak value of the resulting filtered timing jitter plus J_{txout} shall be less than J_{1filt}.

$$H_{jf1}(f) = \frac{jf}{jf + bw1} \qquad f \text{ in Hz}$$

NOTE—j denotes the square root of –1.

When in the normal mode of operation as the SLAVE, receiving valid signals from a compliant PHY operating as the MASTER using the test channel defined in 55.5.1.1, with test channel port A connected to the SLAVE, the peak-to-peak value of the SLAVE TX_TCLK jitter relative to the MASTER TX_TCLK shall be less than J_2 after the receiver is properly receiving the data. When the jitter waveform on TX_TCLK is filtered by a high-pass filter, $H_{jf2}(f)$, having the transfer function below, the peak-to-peak value of the resulting filtered timing jitter plus J_{txout} shall be no more than J_{2filt} greater than the simultaneously measured peak-to-peak value of the MASTER jitter filtered by $H_{jf1}(f)$.

$$H_{jf2}(f) = \frac{jf}{jf + bw2} \qquad f \text{ in Hz}$$

For all high-pass filtered jitter measurements, the peak-to-peak value shall be measured over an unbiased sample of at least 10^6 clock edges. For all unfiltered jitter measurements, the peak-to-peak value shall be measured over an interval of not less than 100 ms and not more than 1 second.

Editor's note: Methodology borrowed from 1000BASE-T. Please suggest alternatives if you want it changed. The parameters J_1, J_2, J_{1filt}, J_{2filt}, bw1, bw2 are TBD.

55.5.5 Transmit clock frequency

The symbol transmission rate on each pair of the master PHY shall be Fs – 0.01%.

55.5.6 Receiver electrical specifications

The PMA shall provide the Receive function specified in 55.4.2.3 in accordance with the electrical specifications of this clause. The patch cabling and interconnecting hardware used in test configurations shall be within the limits specified in 55.7.

55.5.6.1 Receiver differential input signals

Differential signals received at the MDI that were transmitted from a remote transmitter within the specifications of 55.5.3 and have passed through a link specified in 55.7 are received with a BER less than 10^{-12} and sent to the PCS after link reset completion. This specification shall be satisfied by an LDPC code block error rate less than 10^{-TBD}.

55.5.6.2 Receiver frequency tolerance

The receive feature shall properly receive incoming data with a symbol rate within the range Fs – 0.01%.

55.5.7 Common-mode noise rejection

This specification is provided to limit the sensitivity of the PMA receiver to common-mode noise from the cabling system. Common-mode noise generally results when the cabling system is subjected to electromagnetic fields.

The common mode signal that the transceiver shall be subject to, while maintaining link performance, with a LDPC block error rate less than 10^{-TBD}, should be <= 2.8V for f ∈ (1,f_1] MHz, and <= 2.8*f_1/f for f ∈ (f_1, 500) MHz, f_1 in MHz.

Editor's note: Methodology to be agreed upon. Parameter f_1 TBD. Initial recommendation of f_1=80MHz. An appropriate test set-up TBD to realize and test the above specification up to 500MHz, similar to 1000BASE-T standard.

55.5.7.1 Alien Crosstalk noise rejection

While receiving data from a transmitter specified in 55.5.3 through a link segment specified in 55.7 connected to all MDI duplex channels, a receiver shall operate with an LDPC block error rate of less than 10^{-TBD} with four noise sources at the specified levels representing alien NEXT, one connected to each of the four pairs. The noise sources shall be connected to each of the MDI inputs using Category 6 balanced cable of a maximum length of 0.5m. The noise source shall meet the ANEXT specifications in 55.7.3.4.

Editor's note: Please suggest specific ways to generate the ANEXT noise source. Some methods to generate the noise source to model ANEXT were discussed in the presentation:

http://www.ieee802.org/3/an/public/jul04/gupta_1_0704.pdf

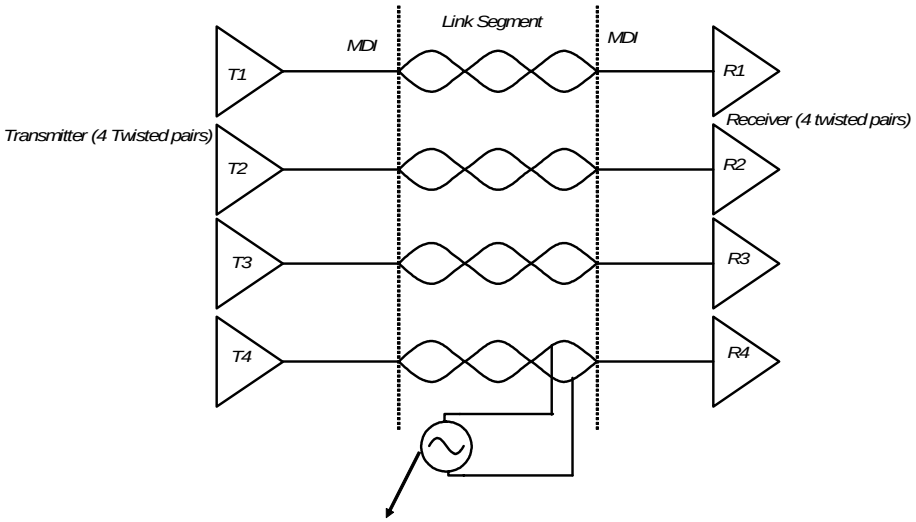


Figure 55–28—ANEXT noise rejection test

55.6 Management interface

10GBASE-T makes extensive use of the management functions provided by the MII Management Interface (see 22.2.4), and the communication and self-configuration functions provided by Auto-Negotiation (Clause 28.)

55.6.1 Support for Auto-Negotiation

All 10GBASE-T PHYs shall provide support for Auto-Negotiation (Clause 28) and shall be capable of operating as MASTER or SLAVE.

Auto-Negotiation is performed as part of the initial set-up of the link, and allows the PHYs at each end to advertise their capabilities (speed, PHY type, half or full duplex) and to automatically select the operating mode for communication on the link. Auto-negotiation signaling is used for the following two primary purposes for 10GBASE-T:

a) To negotiate that the PHY is capable of supporting 10GBASE-T transmission.
b) To determine the MASTER-SLAVE relationship between the PHYs at each end of the link.

This relationship is necessary for establishing the timing control of each PHY. The 1000BASE-T MASTER PHY is clocked from a local source. The SLAVE PHY uses loop timing where the clock is recovered from the received data stream.

55.6.1.1 10GBASE-T use of registers during Auto-Negotiation

A 10GBASE-T PHY shall use the management register definitions and values specified in #CrossRef# Table TBD.

Table 55–6—10GBASE-T Registers

Register	Bit	Name	Description	Type [a]
0	0.15:0	MII control register	Defined in 28.2.4.1.1	RO
1	1.15:0	MII status register	Defined in 28.2.4.1.2	RO
4	4.15:0	Auto-Negotiation advertisement register	The Selector Field (4.4:0) is set to the appropriate code as specified in Annex 28A. The Technology Ability Field bits 4.12:5 are set to the appropriate code as specified in Annexes 28B and 28D. Bit 4.15 is set to logical one to indicate the desired exchange of Next Pages describing the gigabit extended capabilities.	R/W
5	5.15:0	Auto-Negotiation link partner ability register	Defined in 28.2.4.1.4. 10GBASE-T implementations do not use this register to store Auto-Negotiation Link Partner Next Page data.	RO
6	6.15:0	Auto-Negotiation expansion register	Defined in 28.2.4.1.5	RO
7	7.15:0	Auto-Negotiation Next Page transmit register	Defined in 28.2.4.1.6	R/W
8	8.15:0	Auto-Negotiation link partner Next Page register	Defined in 28.2.4.1.8	RO
TBD	TBD	Test mode bits	TBD	R/W

[a] R/W = Read/Write, RO = Read Only, SC = Self Clearing, LH = Latch High

Table 55–6—10GBASE-T Registers *(continued)*

Register	Bit	Name	Description	Type [a]
TBD	TBD	MASTER-SLAVE Manual Config Enable	1=Enable MASTER-SLAVE Manual configuration value 0=Disable MASTER-SLAVE Manual configuration value Default bit value is 0.	R/W
TBD	TBD	MASTER-SLAVE Config Value	1=Configure PHY as MASTER during MASTER-SLAVE negotiation, only when TBD is set to logical one. 0=Configure PHY as SLAVE during MASTER-SLAVE negotiation, only when TBD is set to logical one. Default bit value is 0.	R/W
TBD	TBD	Port type	Bit TBD is to be used to indicate the preference to operate as MASTER (multiport device) or as SLAVE (single-port device) if the MASTER-SLAVE Manual Configuration Enable bit, TBD, is not set. Usage of this bit is described in TBD. 1=Multiport device 0=single-port device	R/W
TBD	TBD	10GBASE-T Full Duplex	1 = Advertise PHY is 10GBASE-T full duplex capable. 0 = Advertise PHY is not 10GBASE-T full duplex capable.	R/W
TBD	TBD	Reserved	Write as 0, ignore on read.	R/W
TBD	TBD	MASTER-SLAVE configuration fault	Configuration fault, as well as the criteria and method of fault detection, is PHY specific. The MASTER-SLAVE Configuration Fault bit will be cleared each time register TBD is read via the management interface and will be cleared by a 10GBASE-T PMA reset. This bit will self clear on Auto-Negotiation enable or Auto-Negotiation complete.This bit will be set if the number of failed MASTER-SLAVE resolutions reaches TBD. For l0GBASE-T, the fault condition will occur when both PHYs are forced to be MASTERs or SLAVEs at the same time using bits TBD and TBD. Bit TBD should be set via the MASTER-SLAVE Configuration Resolution function described in TBD. 1 = MASTER-SLAVE configuration fault detected 0 = No MASTER-SLAVE configuration fault detected	RO/LH/SC
TBD	TBD	MASTER-SLAVE configuration resolution	1 = Local PHY configuration resolved to MASTER 0 = Local PHY configuration resolved to SLAVE	RO
TBD	TBD	Local Receiver Status	1 = Local Receiver OK (loc_rcvr_status=OK) 0 = Local Receiver not OK (loc_rcvr_status=NOT_OK) Defined by the value of loc_rcvr_status as per TBD.	RO
TBD	TBD	Remote Receiver Status	1 = Remote Receiver OK (rem_rcvr_status=OK) 0 = Remote Receiver not OK (rem_rcvr_status=NOT_OK) Defined by the value of rem_rcvr_status as per TBD.	RO
TBD	TBD	LP 10GBASE-T FD	1 = Link Partner is capable of 10GBASE-T full duplex 0 = Link Partner is not capable of 10GBASE-T full duplex This bit is guaranteed to be valid only when the Page received bit (6.1) has been set to 1.	RO

[a] R/W = Read/Write, RO = Read Only, SC = Self Clearing, LH = Latch High

Table 55–6—10GBASE-T Registers *(continued)*

Register	Bit	Name	Description	Type [a]
TBD	TBD	LP 1000T HD	1 = Link Partner is capable of 10GBASE-T half duplex 0 = Link Partner is not capable of 10GBASE-T half duplex This bit is guaranteed to be valid only when the Page received bit (6.1) has been set to 1.	RO
TBD	TBD	Reserved	Reserved	RO
TBD	TBD	Idle Error Count	Bits TBD indicate the Idle Error count, where TBD is the most significant bit. These bits contain a cumulative count of the errors detected when the receiver is receiving idles and PMA_TXMODE.indicate is equal to SEND_N (indicating that both local and remote receiver status have been detected to be OK). The counter is incremented every symbol period that rxerror_status is equal to ERROR. These bits are reset to all zeros when the error count is read by the management function or upon execution of the PCS Reset function and are to be held at all ones in case of overflow (see TBD).	RO/SC
TBD	TBD	Extended status register	See 22.2.4.4	RO

[a] R/W = Read/Write, RO = Read Only, SC = Self Clearing, LH = Latch High

55.6.1.2 10GBASE-T Auto-Negotiation page use

10GBASE-T PHYs shall exchange a 10GBASE-T formatted Next Page, and two 10GBASE-T unformatted Next Pages in sequence, without interruption, as specified in TBD. In the event that the DTE also wishes to advertise 1000BASE-T abilities, the 10GBASE-T pages will follow the 1000BASE-T pages. Additional Next Pages can be exchanged as described in Annex 40C.

Note that the Acknowledge 2 bit is not utilized and has no meaning when used for the 10GBASE-T message page exchange.

Table 55–7—10GBASE-T Base and Next Pages bit assignments

Bit	Bit definition	Register location
BASE PAGE		
D15	1 (to indicate that Next Pages follow)	
D14:D1	As specified in 28.2.1.2	Management register 4
PAGE 0 (Message Next Page)		
M10:M0	9	
PAGE 1 (Unformatted Next Page)		
U10:U4	Reserved transmit as 0	
U3	1000BASE-T full duplex (1 = full duplex and 0 = no full duplex)	GMII register TBD (MASTER-SLAVE Control register)
U2	10GBASE-T port type bit (1 = multiport device and 0 = single-port device)	GMII register TBD (MASTER-SLAVE Control register)
U1	10GBASE-T MASTER-SLAVE Manual Configuration value (1 = MASTER and 0 = SLAVE.) This bit is ignored if TBD =0.	GMII register TBD (MASTER-SLAVE Control register)

Table 55–7—10GBASE-T Base and Next Pages bit assignments *(continued)*

Bit	Bit definition	Register location
U0	10GBASE-T MASTER-SLAVE Manual Configuration Enable (1 = Manual Configuration Enable.) This bit is intended to be used for manual selection in a particular MASTER-SLAVE mode and is to be used in conjunction with bit TBD.	GMII register TBD (MASTER-SLAVE Control register)
PAGE 2 (Unformatted Next Page)		
U10	1000BASE-T MASTER-SLAVE Seed Bit 10 (SB10) (MSB)	MASTER-SLAVE Seed Value (10:0)
U9	1000BASE-T MASTER-SLAVE Seed Bit 9 (SB9)	
U8	1000BASE-T MASTER-SLAVE Seed Bit 8 (SB8)	
U7	1000BASE-T MASTER-SLAVE Seed Bit 7 (SB7)	
U6	1000BASE-T MASTER-SLAVE Seed Bit 6 (SB6)	
U5	1000BASE-T MASTER-SLAVE Seed Bit 5 (SB5)	
U4	1000BASE-T MASTER-SLAVE Seed Bit 4 (SB4)	
U3	1000BASE-T MASTER-SLAVE Seed Bit 3 (SB3)	
U2	1000BASE-T MASTER-SLAVE Seed Bit 2 (SB2)	
U1	1000BASE-T MASTER-SLAVE Seed Bit 1 (SB1)	
U0	1000BASE-T MASTER-SLAVE Seed Bit 0 (SB0)	

55.6.1.3 Sending Next Pages

Implementors who do not wish to send additional Next Pages (i.e., Next Pages in addition to those required to perform PHY configuration as defined in this clause) can use Auto-Negotiation as defined in Clause 28 and the Next Pages defined in TBD. Implementors who wish to send additional Next Pages are advised to consult Annex 40C.

55.6.2 MASTER-SLAVE configuration resolution

Since both PHYs that share a link segment are capable of being MASTER or SLAVE, a prioritization scheme exists to ensure that the correct mode is chosen. The MASTER-SLAVE relationship shall be determined during Auto-Negotiation using TBD with the 10GBASE-T Technology Ability Next Page bit values specified in TBD and information received from the link partner. This process is conducted at the entrance to the FLP LINK GOOD CHECK state shown in the Arbitration state diagram (Figure 28–13.)

The following four equations are used to determine these relationships:

manual_MASTER = U0 * U1
manual_SLAVE = U0 * !U1
single-port device = !U0 * !U2,
multiport device = !U0 * U2

where

U0 is bit 0 of unformatted page 1,
U1 is bit 1 of unformatted page 1, and
U2 is bit 2 of unformatted page 1 (see Table 55–7).

A 10GBASE-T PHY is capable of operating either as the MASTER or SLAVE. In the scenario of a link between a single-port device and a multiport device, the preferred relationship is for the multiport device to be the MASTER PHY and the single-port device to be the SLAVE. However, other topologies may result in contention. The resolution function of TBD is defined to handle any relationship conflicts.

Table 55–8—10GBASE-T MASTER-SLAVE configuration resolution table

Local device type	Remote device type	Local device resolution	Remote device resolution
single-port device	multiport device	SLAVE	MASTER
single-port device	manual_MASTER	SLAVE	MASTER
manual_SLAVE	manual_MASTER	SLAVE	MASTER
manual_SLAVE	multiport device	SLAVE	MASTER
multiport device	manual_MASTER	SLAVE	MASTER
manual_SLAVE	single-port device	SLAVE	MASTER
multiport device	single-port device	MASTER	SLAVE
multiport device	manual_SLAVE	MASTER	SLAVE
manual_MASTER	manual_SLAVE	MASTER	SLAVE
manual_MASTER	single-port device	MASTER	SLAVE
single-port device	manual_SLAVE	MASTER	SLAVE
manual_MASTER	multiport device	MASTER	SLAVE
multiport device	multiport device	The device with the higher SEED value is configured as MASTER, otherwise SLAVE.	The device with the higher SEED value is configured as MASTER, otherwise SLAVE.
single-port device	single-port device	The device with the higher SEED value is configured as MASTER, otherwise SLAVE	The device with the higher SEED value is configured as MASTER, otherwise SLAVE.
manual_SLAVE	manual_SLAVE	MASTER-SLAVE configuration fault	MASTER-SLAVE configuration fault
manual_MASTER	manual_MASTER	MASTER-SLAVE configuration fault	MASTER-SLAVE configuration fault

The rationale for the hierarchy illustrated in Table 55–8 is straightforward. A 10GBASE-T multiport device has higher priority than a single-port device to become the MASTER. In the case where both devices are of the same type, e.g., both devices are multiport devices, the device with the higher MASTER-SLAVE seed bits (SB0...SB10), where SB10 is the MSB, shall become the MASTER and the device with the lower seed value shall become the SLAVE. In case both devices have the same seed value, both should assert link_status_10GigT=FAIL (as defined in 28.3.1) to force a new cycle through Auto-Negotiation. Successful completion of the MASTER-SLAVE resolution shall be treated as MASTER-SLAVE configuration resolution complete.

The method of generating a random or pseudorandom seed is left to the implementor. The generated random seeds should belong to a sequence of independent, identically distributed integer numbers with a uniform distribution in the range of 0 to 2^{11}– 2. The algorithm used to generate the integer should be designed to minimize the correlation between the number generated by any two devices at any given time. A seed

counter shall be provided to track the number of seed attempts. The seed counter shall be set to zero at start-up and shall be incremented each time a seed is generated. When MASTER-SLAVE resolution is complete, the seed counter shall be reset to 0 and bit 10.15 shall be set to logical zero. A MASTER-SLAVE resolution fault shall be declared if resolution is not reached after the generation of seven seeds.

The MASTER-SLAVE Manual Configuration Enable bit (control register bit TBD) and the MASTER-SLAVE Config Value bit (control register bit TBD) are used to manually set a device to become the MASTER or the SLAVE. In case both devices are manually set to become the MASTER or the SLAVE, this condition shall be flagged as a MASTER-SLAVE Configuration fault condition, thus the MASTER-SLAVE Configuration fault bit (status register bit TBD) shall be set to logical one. The MASTER-SLAVE Configuration fault condition shall be treated as MASTER-SLAVE configuration resolution complete and link_status_1GigT shall be set to FAIL, because the MASTER-SLAVE relationship was not resolved. This will force a new cycle through Auto-Negotiation after the link_fail_inhibit_timer has expired. Determination of MASTER-SLAVE values occur on the entrance to the FLP LINK GOOD CHECK state (Figure 28–16) when the highest common denominator (HCD) technology is 10GBASE-T. The resulting MASTER-SLAVE value is used by the 10GBASE-T PHY control (TBD).

If MASTER-SLAVE Manual Configuration is disabled (bit TBD is set to 0) and the local device detects that both the local device and the remote device are of the same type (either multiport device or single-port device) and that both have generated the same random seed, it generates and transmits a new random seed for MASTER-SLAVE negotiation by setting link_status to FAIL and cycling through the Auto-Negotiation process again.

The MASTER-SLAVE configuration process returns one of the three following outcomes:

a) *Successful:* Bit TBD of the 10GBASE-T Status Register is set to logical zero and bit TBD is set to logical one for MASTER resolution or for logical zero for SLAVE resolution. 10GBASE-T returns control to Auto_Negotiation (at the entrance to the FLP LINK GOOD CHECK state in Figure 28–16) and passes the value MASTER or SLAVE to PMA_CONFIG.indicate (see TBD.)
b) *Unsuccessful:* link_status_10GigT is set to FAIL and Auto-Negotiation restarts (see Figure 28–16.)
c) *Fault detected:* (This happens when both end stations are set for manual configuration and both are set to MASTER or both are set to SLAVE.) Bit TBD of the 10GBASE-T Status Register is set to logical one to indicate that a configuration fault has been detected. This bit also is set when seven attempts to configure the MASTER SLAVE relationship via the seed method have failed. When a fault is detected, link_status_10GigT is set to FAIL, causing Auto-Negotiation to cycle through again.

NOTE—MASTER-SLAVE arbitration only occurs if 10GBASE-T is selected as the highest common denominator; otherwise, it is assumed to have passed this condition

55.7 Link segment characteristics

10GBASE-T is designed to operate over ISO/IEC 11801 Class E or Class F 4-Pair balanced cabling that meets the additional requirements specified in 55.7. Each of the four pairs supports an effective data rate of (2500) Mbps in each direction simultaneously. The term “link segment” used in this clause refers to four duplex channels. The term “duplex channel” will be used to refer to a single channel with full duplex capability. Specifications for a link segment apply equally to each of the four duplex channels. All implementations of the balanced cabling link segment specification shall be compatible at the MDI.

55.7.1 Cabling system characteristics

The cabling system used to support 10GBASE-T requires 4 pairs of ISO/IEC 11801 Class E or Class F balanced cabling with a nominal impedance of 100 W.

Additionally:

a) 10GBASE-T uses a star topology with Class E or Class F balanced cabling used to connect PHY entities.
b) 10GBASE-T is an ISO/IEC 11801 Class E and Class F application with the additional transmission requirements specified in 55.7.

55.7.2 Link transmission parameters

The transmission parameters contained in this subclause are specified to ensure that a Class E link segment of at least 55 to 100 meters and a Class F link segment of at least 100 meters will provide a reliable medium. The transmission parameters of the link segment include insertion loss, delay parameters, characteristic impedance, NEXT loss, ELFEXT loss, and return loss. In addition, link segment requirements are specified in subclause 55.7.3.2 for alien crosstalk.

Link segment testing shall be conducted using source and load impedances of 100 W.

The link segment transmission parameters of insertion loss and ELFEXT loss specified in 55.7 are ISO/IEC 11801 Class E specifications extended by extrapolating the formulas to a frequency up to 500 MHz. The link segment transmission parameters of NEXT loss, MDNEXT loss and Return Loss specified in 55.7 are ISO/IEC 11801 Class E specifications extended beyond 250 MHz by utilizing the equations referenced in TIA TR42 D1.0 TSB-155.

Note—The Class F and Augmented Category 6 (Augmented Class E) channel limits meet or exceed the Clause 55 Class E link segment specifications. Class F and Augmented Category 6 (Augmented Class E) specifications are referenced only when applicable to requirements specific to Class F and Augmented Category 6 such as power sum alien NEXT (PS ANEXT) and Insertion Loss.

55.7.2.1 Insertion loss

The insertion loss of each duplex channel shall be less than

$$1.05\left(1.82 \cdot \sqrt{f} + 0.169 \cdot f + \frac{0.25}{\sqrt{f}}\right) + 4 \cdot 0.02 \cdot \sqrt{f} \qquad \text{(dB)}$$

at all frequencies from 1 MHz to 500 MHz. This includes the attenuation of the balanced cabling pairs, including work area and equipment cables plus connector losses within each duplex channel. The insertion loss specification shall be met when the duplex channel is terminated in 100 W.

55.7.2.2 Differential characteristic impedance

The nominal differential characteristic impedance of each link segment duplex channel, which includes cable cords and connecting hardware, is 100 W for all frequencies between 1 MHz and 500 MHz.

55.7.2.3 Return loss

Each link segment duplex channel shall meet or exceed the return loss specified in the following equation at all frequencies from 1 MHz to 500 MHz. The reference impedance shall be 100 W .

$$\text{Return\_Loss}(f) = \begin{cases} 19 & 1 \le f < 10 \\ 24 - 5\log 10(f) & 10 \le f < 40 \\ 32 - 10\log 10(f) & 40 \le f < 400 \\ 6 & 400 \le f \le 500 \end{cases} \quad \text{(dB)}$$

where f is the frequency in MHz.

55.7.3 Coupling parameters

55.7.3.1 Coupling parameters between duplex channels

In order to limit the noise coupled into a duplex channel from adjacent duplex channels, Near-End Crosstalk (NEXT) loss and Equal Level Far-End Crosstalk (ELFEXT) loss are specified for each link segment. In addition, each duplex channel can be disturbed by more than one duplex channel. To ensure the total NEXT loss and FEXT loss coupled into a duplex channel is limited, multiple disturber Near-End Crosstalk (MDNEXT) and multiple disturber ELFEXT (MDELFEXT) loss is specified.

55.7.3.1.1 Near-End Crosstalk (NEXT)

55.7.3.1.1.1 Differential Near-End Crosstalk

In order to limit the crosstalk at the near end of a link segment, the differential pair-to-pair Near-End Crosstalk (NEXT) loss between a duplex channel and the other three duplex channels is specified to meet the bit error rate objective specified in 55.1. The NEXT loss between any two duplex channels of a link segment shall be at least

$$-20 \cdot \log 10\left(10^{\frac{74.3 - 15\log 10(f)}{-20}} + 2 \cdot 10^{\frac{94 - 20\log 10(f)}{-20}}\right) \quad \text{(dB)}$$

where f is the frequency $(1 \le f < 330)$ in MHz.

The NEXT loss between any two duplex channels of a link segment shall be at least

$$31 - 50 \cdot \log 10\left(\frac{f}{330}\right) \quad \text{(dB)}$$

where f is the frequency $(330 \le f \le 500)$ in MHz.

55.7.3.1.1.2 Multiple Disturber Near-End Crosstalk (MDNEXT) loss

Since four duplex channels are used to transfer data between PMDs, the NEXT that is coupled into a data carrying channel will be from the three adjacent disturbing duplex channels.

To ensure the total NEXT coupled into a duplex channel is limited, multiple disturber NEXT loss is specified as the power sum of the individual NEXT losses. The Power Sum loss between a duplex channel and the three adjacent disturbers shall be greater than

$$-20 \cdot \log 10\left(10^{\frac{72.3 - 15\log 10(f)}{-20}} + 2 \cdot 10^{\frac{90 - 20\log 10(f)}{-20}}\right) \quad \text{(dB)}$$

where f is the frequency $1 \le f < 330$ in MHz and

$$28 - 42 \cdot \log 10\left(\frac{f}{330}\right) \quad \text{(dB)}$$

where f is the frequency $(330 \le f \le 500)$ in MHz.

55.7.3.1.1.3 Multiple-Disturber Power Sum Near-End Crosstalk (PS NEXT) loss

PS NEXT loss is determined by summing the power of the three individual pair-to-pair differential NEXT loss values over the frequency range 1 MHz to 500 MHz. as follows:

$$-10 \cdot \log 10 \sum_{i=1}^{n} 10^{\frac{-NL(f)i}{10}} \quad \text{(dB)}$$

where

NL(f)i is the magnitude in dB of NEXT loss at frequency f of pair combination i

i is the 1, 2, or 3 (pair-to-pair combination)

n is the number of pair-to-pair combinations

55.7.3.1.2 Far-End Crosstalk (FEXT)

55.7.3.1.2.1 Equal Level Far-End Crosstalk (ELFEXT) loss

Equal Level Far-End Crosstalk (ELFEXT) loss is specified in order to limit the crosstalk at the far end of each link segment duplex channel and meet the BER objective specified in 55.1.1. Far-End Crosstalk

(FEXT) is crosstalk that appears at the far end of a duplex channel (disturbed channel), which is coupled from another duplex channel (disturbing channel) with the noise source (transmitters) at the near end.

Editor's Note: For 1000BASE-T the error rate is specified as symbol error rate, frame error rate and bit error rate. For 10GBASE-T D1.0, as a starting point, the BER objective of 10^-12 specified in 55.1 will be utilized throughout subclause 55.7.

FEXT loss is defined as

$$\text{FEXT\_Loss(f)} = 20 \cdot \log 10\left(\frac{V_{pds}(f)}{V_{pcn}(f)}\right) \qquad \text{(dB)}$$

and ELFEXT_Loss is defined as

$$\text{ELFEXT\_Loss(f)} = 20 \cdot \log 10\left(\frac{V_{pds}(f)}{V_{pcn}(f)}\right) - \text{SLS\_Loss(f)} \qquad (dB)$$

where

*V*pds	is the peak voltage of disturbing signal (near-end transmitter)
*V*pcn	is the peak crosstalk noise at far end of disturbed channel
SLS_Loss	is the insertion loss of disturbed channel in dB

The worst pair ELFEXT loss between any two duplex channels shall be greater than

$$-20 \cdot \log 10\left(10^{\frac{67.8 - 20\log 10(f)}{-20}} + 4 \cdot 10^{\frac{83.1 - 20\log 10(f)}{-20}}\right) \qquad \text{(dB)}$$

where *f* is the frequency over the range of 1 MHz to 500 MHz.

55.7.3.1.2.2 Multiple Disturber Equal Level Far-End Crosstalk (MDELFEXT) loss

Since four duplex channels are used to transfer data between PMDs, the FEXT that is coupled into a data carrying channel will be from the three adjacent disturbing duplex channels. To ensure the total FEXT coupled into a duplex channel is limited, multiple disturber ELFEXT loss is specified as the power sum of the individual ELFEXT losses. The Power Sum loss between a duplex channel and the three adjacent disturbers shall be greater than

$$-20 \cdot \log 10\left(10^{\frac{64.8 - 20\log 10(f)}{-20}} + 4 \cdot 10^{\frac{80.1 - 20\log 10(f)}{-20}}\right) \quad \text{(dB)}$$

where f is the frequency over the range of 1 MHz to 500 MHz.

55.7.3.1.2.3 Multiple-Disturber Power Sum Equal Level Far-End Crosstalk (PS ELFEXT) loss

PS ELFEXT loss is determined by summing the power of the three individual pair-to-pair differential ELFEXT loss values over the frequency range 1 MHz to 500 MHz as follows:

$$\text{PSELFEXT\_Loss(f)} = -10 \cdot \log_{10} \sum_{i=1}^{i=n} 10^{\frac{-EL(f)i}{10}}$$

where

EL(f)i is the magnitude of ELFEXT loss at frequency f of pair combination i
i is the 1, 2, or 3 (pair-to-pair combination)
n is the number of pair-to-pair combinations

55.7.3.2 Coupling parameters between link segments

Noise coupled between the disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments is referred to as alien crosstalk noise.

Editor's Note: Text needs to be added to clearly identify the alien crosstalk dependencies.

55.7.3.2.1 Multiple Disturber Alien Near-End Crosstalk (MDANEXT) loss

In order to limit the alien crosstalk at the near end of a link segment, the differential pair-to-pair Near-End Crosstalk (NEXT) loss between the disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments is specified to meet the bit error rate objective specified in 55.1. To ensure the total Alien NEXT coupled into a duplex channel is limited, multiple disturber Alien NEXT loss is specified as the power sum of the individual Alien NEXT disturbers.

55.7.3.2.1.1 Multiple-Disturber Power Sum Near-End Crosstalk (PS ANEXT) loss

ANEXT loss is determined by summing the power of the individual pair-to-pair differential Alien NEXT loss values over the frequency range 1 MHz to 500 MHz. as follows:

$$-10 \cdot \log 10 \sum_{i=1}^{n} 10^{\frac{-AN(f)i}{10}} \quad \text{(dB)}$$

where

AN(f)i is the magnitude in dB of ANEXT loss at frequency f of pair combination i

i is the pair-to-pair combination (1 to n)

n is the number of pair-to-pair combinations between link segments

The Power Sum ANEXT loss between a disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments is defined by the equations:

$$PSANEXT \geq \begin{cases} X1 - 10\log 10(\frac{f}{100}) \quad (dB) & 1 \leq f \leq 100 \\ X1 - 15\log 10(\frac{f}{100}) \quad (dB) & 100 < f \leq 500 \end{cases}$$

where f is the frequency in MHz and X1 = the intercept at f=100 MHz. The intercept is referred to as the PS ANEXT constant.

55.7.3.2.1.2 PS ANEXT for a Class E Channel

For a 100 meter Class E channel with the maximum insertion loss of 55.7.2.1 the PS ANEXT loss between a disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments shall be greater than

$$PSANEXT \geq \begin{cases} 62 - 10\log 10(\frac{f}{100}) \quad (dB) & 1 \leq f \leq 100 \\ 62 - 15\log 10(\frac{f}{100}) \quad (dB) & 100 < f \leq 500 \end{cases}$$

55.7.3.2.1.3 PS ANEXT for a Class F Channel

For a 100 meter Class F channel the PS ANEXT loss between a disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments shall be greater than

$$PSANEXT \geq \begin{cases} 60 - 10\log 10(\frac{f}{100}) \quad (dB) & 1 \leq f \leq 100 \\ 60 - 15\log 10(\frac{f}{100}) \quad (dB) & 100 < f \leq 500 \end{cases}$$

where f is the frequency in MHz.

The PS ANEXT for a Class F channel specified in 55.7.3.2.1.3 assumes the maximum insertion loss of a Class F channel in 55.7.3.3.1

Editors Note: Alien crosstalk is not adequately specified in the ISO/IEC 11801 or TIA cabling standards. The PS ANEXT limits for both Class F and Class E are the minimum requirements for 100 meter operation and are not intended to represent the PS ANEXT performance limits of the cabling (i.e., the PS ANEXT performance of the cabling may be better than the minimum requirements specified in 10GBASE-T). TIA TR42 has initiated Project SP-3-4426-AD10 to develop augmented Category 6 cabling. The resulting requirements will be presented in a new revision or addendum to the TIA-568-B standard.

55.7.3.2.1.4 PS ANEXT for an Augmented Category 6 (Augmented Class E) Channel

For a 100 meter Augmented Category 6 (Augmented Class E) channel the PS ANEXT loss between a disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments shall be greater than

$$PSANEXT \ge \begin{cases} 60 - 10\log 10(\frac{f}{100}) \quad (dB) & 1 \le f \le 100 \\ 60 - 15\log 10(\frac{f}{100}) \quad (dB) & 100 < f \le 500 \end{cases}$$

The PS ANEXT for an Augmented Category 6 (Augmented Class E) Channel specified in 55.7.3.2.1.4 assumes the maximum insertion loss of an Augmented Category 6 (Augmented Class E) channel in 55.7.3.3.2.

55.7.3.2.2 Multiple Disturber Alien Far-End Crosstalk (MDAFEXT) loss (ffs)

55.7.3.2.2.1 Multiple -Disturber Power Sum Alien Far-End Crosstalk (PS AFEXT) loss (ffs)

55.7.3.3 PS ANEXT loss to insertion loss ratio requirements

To ensure reliable operation, a minimum signal to noise ratio (SNR) must be maintained. The minimum SNR is assured for 100 meters of Class E by meeting the requirements of 55.7.1 through 55.7.3.2.1.2 and for 100 meters of Class F by meeting the additional requirement of 55.7.3.2.1.3.

The PS ANEXT loss requirement of 55.7.3.2.1.2 can be relaxed based on a reduction in the maximum insertion loss specified in 55.7.2.1. The insertion loss reduction can be achieved by scaling the length of the Class E link segment or using Class F cabling for the link segment with the additional requirement of 55.7.3.2.1.3 or using Augmented Category 6 (Augmented Class E) Channel for the link segment as specified in 55.7.3.2.1.4 and 55.7.3.3.2.

55.7.3.3.1 Insertion Loss for a Class F Channel

The PS ANEXT for a Class F Channel assumes the maximum insertion loss of a Class F channel.

The insertion loss of a Class F duplex channel shall be less than

$$1.05\left(1.8 \cdot \sqrt{f} + 0.01 \cdot f + \frac{0.2}{\sqrt{f}}\right) + 4 \cdot 0.02 \cdot \sqrt{f} \qquad \text{(dB)}$$

at all frequencies from 1 MHz to 500 MHz. This includes the attenuation of the balanced cabling pairs, including the cords and connecting hardware losses within each duplex channel. The insertion loss specification shall be met when the duplex channel is terminated in 100 W.

NOTE— The Class F insertion loss is an improvement of 2.1 dB at 250 MHz over the Class E insertion loss specifications resulting in a 2 dB relaxation in the Class F PS ANEXT requirement.

55.7.3.3.2 Insertion Loss for an Augmented Category 6 (Augmented Class E) Channel

The insertion loss of an Augmented Category 6 (Augmented Class E) Channel shall be less than

$$1.05\left(1.8 \cdot \sqrt{f} + 0.01 \cdot f + \frac{0.2}{\sqrt{f}}\right) + 4 \cdot 0.02 \cdot \sqrt{f} \qquad \text{(dB)}$$

at all frequencies from 1 MHz to 500 MHz. This includes the attenuation of the balanced cabling pairs, including the cords and connecting hardware losses within each duplex channel. The insertion loss specification shall be met when the duplex channel is terminated in 100 W.

55.7.3.3.3 Insertion Loss Scaling

For the purpose of adjusting the PS ANEXT the insertion loss is assumed to scale linearly with length.

The Scaled Class E IL is defined by the following equation:

$$\text{Scaled\_Class\_E\_IL} = \frac{\text{Length\_m}}{100} \cdot 1.05\left(1.82 \cdot \sqrt{f} + 0.0169 \cdot f + \frac{0.25}{\sqrt{f}}\right) + 4 \cdot 0.02 \cdot \sqrt{f} \text{ (dB)}$$

where Length is in meters.

55.7.3.3.4 Insertion Loss of a Category 6 channel of 55 meters

ISO/IEC 11801 classes for balanced cabling refer to cabling channel distances of 100 meters. For cabling channels less than 100 meters the Category of the components comprising the channel applies (e.g., Category 6 components provide Class E balanced cabling performance).

The insertion loss of a Category 6 channel of 55 meters is defined by the following equation:

$$\text{Scaled\_Class\_E\_IL}(55m) = \frac{55}{100} \cdot 1.05\left(1.82 \cdot \sqrt{f} + 0.0169 \cdot f + \frac{0.25}{\sqrt{f}}\right) + 4 \cdot 0.02 \cdot \sqrt{f} \text{ (dB)}$$

55.7.3.4 PS ANEXT Adjustment

The adjusted PS ANEXT loss requirement is determined by first calculating the PS ANEXT_constant and utilizing the constant in the PS ANEXT limit line model.

The PS ANEXT_constant is defined by the following equation:

$$\text{PSANEXT\_constant} = 62 - (\text{CE\_IL\_250}MHz - \text{SCE\_IL\_250}MHz) \cdot \frac{15}{15.6} \qquad \text{(dB)}$$

where

CE_IL_250 MHz is the Class E insertion loss at 250 MHz

SCE_IL_250 MHz is the scaled Class E insertion loss at 250 MHz.

55.7.3.4.1 PS ANEXT for a Category 6 channel of 55 meters

For a Category 6 channel of 55 meters with worst case insertion loss of 55.7.3.3.4 the PS ANEXT loss between a disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments shall be greater than

$$PSANEXT > \left\{ \begin{array}{ll} 47 - 10\log 10(\frac{f}{100}) \quad (dB) & 1 \leq f \leq 100 \\ 47 - 15\log 10(\frac{f}{100}) \quad (dB) & 100 < f \leq 500 \end{array} \right\}$$

where f is the frequency in MHz.

55.7.4 Delay

In order to simultaneously send data over four duplex channels in parallel, the propagation delay of each duplex channel as well as the difference in delay between any two of the four channels are specified. This ensures the 10 Gbps data that is divided across four channels can be properly reassembled at the far-end receiver.

55.7.4.1 Maximum link delay

The propagation delay of a link segment shall not exceed 570 ns at all frequencies between 2 MHz and 500 MHz.

55.7.4.2 Link delay skew

The difference in propagation delay, or skew, between all duplex channel pair combinations of a link segment, under all conditions, shall not exceed 50 ns at all frequencies from 2 MHz to 500 MHz. It is a further functional requirement that, once installed, the skew between any two of the four duplex channels due to environmental conditions shall not vary more than 10 ns within the above requirement.

55.7.5 Noise environment

Editor's Note: The noise environment (55.7.5) sub clause is extracted from 1000BASE-T specification with minor changes. This text will likely evolve to reflect the 10GBASE-T noise environment assumptions.

The 10GBASE-T noise environment consists of noise from many sources. The primary noise sources that impact the objective BER are NEXT and echo interference, which are reduced to a small residual noise using cancellers. The remaining noise sources, which are secondary sources, are discussed in the following list. The 10GBASE-T noise environment consists of the following:

a) Echo from the local transmitter on the same duplex channel (cable pair). Echo is caused by the hybrid function used to achieve simultaneous bi-directional transmission of data and by impedance mismatches in the link segment. It is impractical to achieve the objective BER without using echo cancellation. Since the symbols transmitted by the local disturbing transmitter are available to the cancellation processor, echo interference can be reduced to a small residual noise using echo cancellation methods.

b) Near-End Crosstalk (NEXT) interference from the local transmitters on the duplex channels (cable pairs) of the link segment. Each receiver will experience NEXT interference from three adjacent

transmitters. NEXT cancellers are used to reduce the interference from each of the three disturbing transmitters to a small residual noise. NEXT cancellation is possible since the symbols transmitted by the three disturbing local transmitters are available to the cancellation processor.

c) Far-End Crosstalk (FEXT) noise at a receiver is from three disturbing transmitters at the far end of the duplex channel (cable pairs) of the link segment. FEXT noise can be cancelled in the same way as echo and NEXT interference although the symbols from the remote transmitters are not immediately available.

d) Inter-Symbol Interference (ISI) noise. ISI is the extraneous energy from one signaling symbol that interferes with the reception of another symbol on the same channel.

e) Noise from non-idealities in the duplex channel, transmitters, and receivers; for example, DAC/ADC non-linearity, electrical noise (shot and thermal), and non-linear channel characteristics.

f) Noise coupled between link segments. Noise coupled between the disturbed duplex channel in a link segment and the disturbing duplex channels in other link segments is referred to as alien crosstalk noise.Since the transmitted symbols from the alien NEXT noise source are not available to the cancellation processor (they are in another cable), it is very difficult to cancel the alien NEXT noise. To ensure robust operation the alien NEXT noise must meet the specification of 55.7.5.

g) The background noise for 10GBASE-T is expected not to exceed -150 dBm/Hz.. A background noise limit of -150 dBm/Hz was assumed in the 10GBASE-T Matlab simulation models.

55.8 MDI specification

This subclause defines the MDI. The link topology requires a crossover function in a DTE-to-DTE connection. See 55.4.4 for a description of the automatic MDI/MDI-X configuration.

55.8.1 MDI connectors

Eight-pin connectors meeting the requirements of subclause 3 and Figures 1 through 4 of IEC 60603-7: 1995 shall be used as the mechanical interface to the balanced cabling. The plug connector shall be used on the balanced cabling and the jack on the PHY. These connectors are depicted (for informational use only) in Figure 55–29 and Figure 55–30. The assignment of PMA signals to connector contacts for PHYs is shown in Table 55–9. The PHY shall be capable of reversing the polarity of the contacts for any PMA signal to correct for a mis-wired channel between any two PHY entities within the link segment.

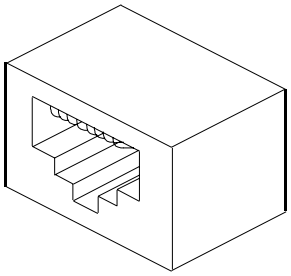


Figure 55–29—MDI connector

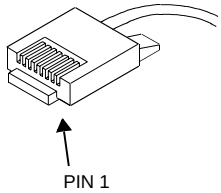


Figure 55–30—Balanced cabling connector

Table 55–9—Assignment of PMA signal to MDI and MDI-X pin-outs

Contact	MDI	MDI-X
1	BI_DA+	BI_DB+
2	BI_DA–	BI_DB–
3	BI_DB+	BI_DA+
4	BI_DC+	BI_DD+
5	BI_DC–	BI_DD–
6	BI_DB–	BI_DA–
7	BI_DD+	BI_DC+
8	BI_DD–	BI_DC–

55.8.2 Crossover function

It is a functional requirement that a crossover function be implemented in every link segment to support the operation of Auto-Negotiation. The crossover function connects the transmitters of one PHY to the receivers of the PHY at the other end of the link segment. Crossover functions may be implemented internally to a PHY or else-where in the link segment. The MDI connector for a 10GBASE-T PHY shall be marked with the graphical symbol X. The crossover function specified here is not compatible with the crossover function specified in 14.5.2 for pairs TD and RD.

When a link segment connects a single-port device to a multiport device, it is recommended that the crossover be implemented in the PHY local to the multiport device. It is recommended that the crossover be visible to an installer from one of the PHYs. It is further recommended that, in networks in which the topology identifies either a central backbone segment or a central device, the PHY furthest from the central element be assigned the external crossover to maintain consistency.

Implicit implementation of the crossover function within a twisted-pair cable or at a wiring panel, while not expressly forbidden, is beyond the scope of this standard.

Editor's note: The MDI crossover function has been made mandatory rather than optional.

55.8.3 MDI electrical specifications

The MDI connector (jack) when mated with a specified balanced cabling connector (plug) shall meet the electrical requirements for Category 6 connecting hardware for use with 100-ohm Category 6 cable as specified in ANSI/TIA/EIA-568-B.2:2002 and ISO/IEC 11801:2002.

The mated MDI/balanced cabling connector pair shall have a FEXT loss not less than $43.1 - 20\log_{10}(f/100)$ (where f is the frequency over the range 1 MHz to 500 MHz) between all contact pair combinations shown in Table 55–9.

No spurious signals shall be emitted onto the MDI when the PHY is held in power-down mode (as defined in 22.2.4.1.5) independent of the value of TX_EN, when released from power-down mode, or when external power is first applied to the PHY.

55.8.3.1 MDI return loss

The differential impedance at the MDI for each transmit/receive channel shall be such that any reflection due to differential signals incident upon the MDI from a balanced cabling having an impedance of 100 W – 15% is attenuated, relative to the incident signal, at least 16 dB over the frequency range of 1.0 MHz to 40 MHz and at least $16 - 20\log_{10}(f/40)$ dB over the frequency range 40 MHz to 500 MHz (f in MHz). This return loss shall be maintained at all times when the PHY is transmitting data or control symbols.

55.8.3.2 MDI impedance balance

Impedance balance is a measurement of the impedance-to-ground difference between the two MDI contacts used by a duplex link channel and is referred to as common-mode-to-differential-mode impedance balance. Over the frequency range 1.0 MHz to 500.0 MHz, the common-mode-to-differential-mode impedance balance of each channel of the MDI shall exceed

$$45 - 4\log_{10}\left(\frac{f}{50}\right) \quad \text{dB}$$

where f is the frequency in MHz when the transmitter is transmitting random or pseudo random data. Test-mode 4 may be used to generate an appropriate transmitter output.

The balance is defined as

$$20\log_{10}\left(\frac{E_{cm}}{E_{dif}}\right)$$

where E_{cm} is an externally applied sine wave voltage as shown in Figure 55–31 and E_{dif} is the resulting waveform due only to the applied sine wave and not the transmitted data.

NOTES

1—Triggered averaging can be used to separate the component due to the applied common-mode sine wave from the transmitted data component.

2—The imbalance of the test equipment (such as the matching of the test resistors) must be insignificant relative to the balance requirements.

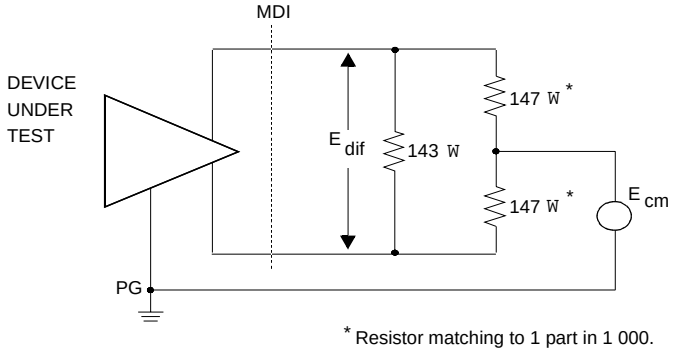


Figure 55–31—MDI impedance balance test circuit

Editor's note: The impedance balance equation was obtained by processing the data provided by Pulse on the magnetics. Other factors contributing to the imbalance may require us to modify the above requirement. Please provide feedback on the feasibility of the above proposal.

55.8.3.3 MDI common-mode output voltage

The magnitude of the total common-mode output voltage, E_{cm_out}, on any transmit circuit, when measured as shown in Figure 55–32, shall be less than 15 mV peak-to-peak when transmitting data.

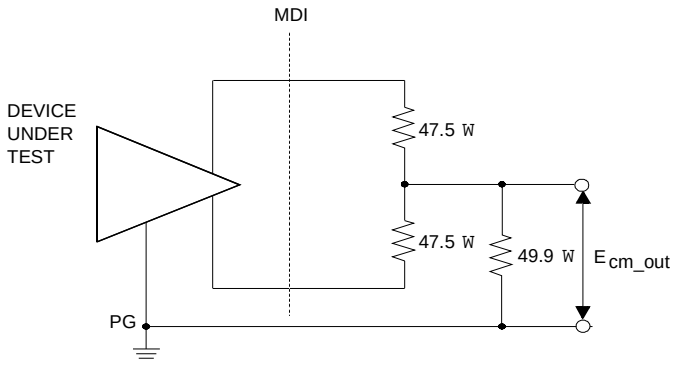


Figure 55–32—Common-mode output voltage test circuit

Editor's note: 1000BASE-T specified 50 mV as the maximum permissible common-mode output voltage. That number has been reduced to 15 mV to because reducing it should make it easier to pass EMI. We need feedback from the Task Force on whether this value is feasible from an implementation point of view from both the PHY vendors and the magnetics manufacturers.

NOTE—The imbalance of the test equipment (such as the matching of the test resistors) must be insignificant relative to the balance requirements.

55.8.3.4 MDI fault tolerance

Each wire pair of the MDI shall, under all operating conditions, withstand without damage the application of short circuits of any wire to any other wire within the 4-pair cable for an indefinite period of time and shall resume normal operation after the short circuit(s) are removed. The magnitude of the current through such a short circuit shall not exceed 300 mA.

Each wire pair shall withstand without damage a 1000 V common-mode impulse applied at E_{cm} of either polarity (as indicated in Figure 55–33). The shape of the impulse shall be 0.3/50 ms (300 ns virtual front time, 50 ms virtual time of half value), as defined in IEC 60060.

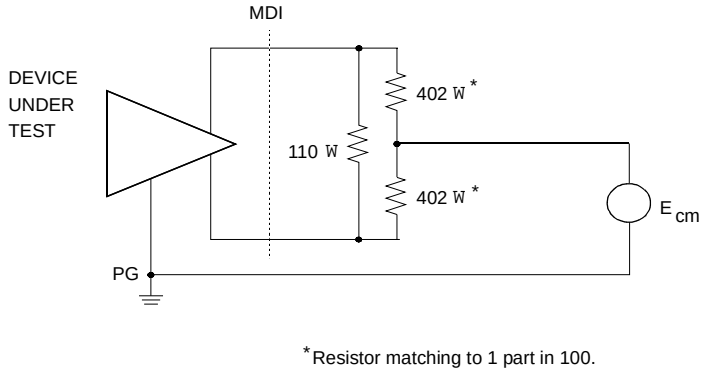


Figure 55–33—MDI fault tolerance test circuit

55.9 Environmental specifications

55.9.1 General safety

All equipment meeting this standard shall conform to IEC 60950-1: 2001.

55.9.2 Network safety

This subclause sets forth a number of recommendations and guidelines related to safety concerns; the list is neither complete nor does it address all possible safety issues. The designer is urged to consult the relevant local, national, and international safety regulations to ensure compliance with the appropriate requirements.

LAN cabling systems described in this subclause are subject to at least four direct electrical safety hazards during their installation and use. These hazards are as follows:

a) Direct contact between LAN components and power, lighting, or communications circuits.
b) Static charge buildup on LAN cabling and components.
c) High-energy transients coupled onto the LAN cabling system.
d) Voltage potential differences between safety grounds to which various LAN components are connected.

Such electrical safety hazards must be avoided or appropriately protected against for proper network installation and performance. In addition to provisions for proper handling of these conditions in an operational

system, special measures must be taken to ensure that the intended safety features are not negated during installation of a new network or during modification or maintenance of an existing network.

Editor's note: References have been updated; the frequency range of interest is now from 1MHz to 500MHz. The equations have been updated to reflect this and the formulae have been adjusted.

55.9.3 Installation and maintenance guidelines

It is a mandatory requirement that sound installation practice, as defined by applicable local codes and regulations, be followed in every instance in which such practice is applicable.

It is a mandatory requirement that, during installation of the cabling plant, care be taken to ensure that non-insulated network cabling conductors do not make electrical contact with unintended conductors or ground.

55.9.4 Telephone voltages

The use of building wiring brings with it the possibility of wiring errors that may connect telephony voltages to 10GBASE-T. Other than voice signals (which are low voltage), the primary voltages that may be encountered are the "battery" and ringing voltages. Although there is no universal standard. The following maximums generally apply.

a) Battery voltage to a telephone line is generally 56 Vdc applied to the line through a balanced 400 W source impedance.
a) Ringing voltage is a composite signal consisting of an ac component and a dc component. The ac component is up to 175 V peak at 20 Hz to 60Hz with a 100 W source resistance. The dc component is 56 Vdc with a 300 to 600 W source resistance. Large reactive transients can occur at the start and end of each ring interval.

Although 10GBASE-T equipment is not required to survive such wiring hazards without damage, application of any of the above voltage shall not result in any safety hazard.

55.9.5 Electromagnetic emissions

A system integrating the 10GBASE-T shall comply with applicable local and national codes for the limitation of electromagnetic interference.

55.9.6 Temperature and humidity

A system integrating the 10GBASE-T is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration). Specific requirements and values for these parameters are considered to be beyond the scope of this standard.

55.10 PHY labeling

It is recommended that each PHY (and supporting documentation) be labeled in a manner visible to the user with at least the following parameters:

a) Data rate capability in Mb/s
b) Power level in terms of maximum current drain (for external PHYs)
c) Port type (i.e., 10GBASE-T)
d) Any applicable safety warnings

55.11 Delay constraints

In full duplex mode, predictable operation of the MAC Control PAUSE operation (Clause 31, Annex 31B) also demands that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sublayer, and PHY implementors must conform to certain delay maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices.

The reference point for all MDI measurements is the peak point of the mid-cell transition corresponding to the reference code-bit, as measured at the MDI.

55.11.1 MDI to XGMII delay constraints

Every 10GBASE-T PHY associated with a XGMII shall comply with the bit delay constraints specified in Table 55–10 for full duplex operation. These constraints apply for all 10GBASE-T PHYs.

Table 55–10—MDI to GMII delay constraints (full duplex mode)

Sublayer measurement points	Event	Min (bit times)	Max (bit times)	Input timing reference	Output timing reference
XGMII MDI XGMII MDI	SFD coming in on XGMII on one PHY and coming out of the XGMII on the other PHY with two PHYs connected back to back with 10m cable	—	TBD	TBD	TBD
XGMII MDI	TBD	—	TBD	TBD	TBD

Editor's note: Delay is measureable easily on two PHYs connected back to back however this delay will vary depending on timing of arrival of the SFD on the XGMII relative to start of LDPC code. Delay from XGMII to MDI will be hard to measure because all symbols are encoded and buried inside LDPC code words. It may make sense to have an optional flag put in which is detectable on the MDI port solely to enable easy detection of delays while testing.

55.12 Protocol implementation conformance statement (PICS) proforma for Clause 55—Physical coding sublayer (PCS), physical medium attachment (PMA) sublayer and baseband medium, type 10GBASE-T[3]

The supplier of a protocol implementation that is claimed to conform to this clause shall complete the Protocol Implementation Conformance Statement (PICS) proforma listed in the following subclauses.

Instructions for interpreting and filling out the PICS proforma may be found in Clause 21.

[3]*Copyright release for PICS proformas:* Users of this standard may freely reproduce the PICS proforma in this subclause so it can be used for its intended purpose and may further publish the completed PICS.

55.12.1 Identification

55.12.1.1 Implementation identification

Supplier	
Contact point for queries about the PICS	
Implementation Name(s) and Version(s)	
Other information necessary for full identification—e.g., name(s) and version(s) for machines and/or operating systems; System Name(s)	
NOTES 1—Only the first three items are required for all implementations; other information may be completed as appropriate in meeting the requirements for the identification. 2—The terms Name and Version should be interpreted appropriately to correspond with a supplier's terminology (e.g., Type, Series, Model).	

55.12.1.2 Protocol summary

Identification of protocol specification	IEEE Std 802.an:?, Clause 55, Physical coding sublayer (PCS), physical medium attachment (PMA) sublayer, and baseband medium, type 10GBASE-T
Identification of amendments and corrigenda to this PICS proforma which have been completed as part of this PICS	
Have any Exceptions items been required? No [ ] Yes [ ] (See Clause 21—The answer Yes means that the implementation does not conform to the standard)	
Date of Statement	

55.12.2 Major capabilities/options

Item	Feature	Subclause	Status	Support	Value/Comment
*XGMII	PHY associated with XGMII	TBD	O	Yes [ ] No [ ]	

55.12.3 Clause conventions

Item	Feature	Subclause	Status	Support	Value/Comment
CCO1	The values of all components in test circuits shall be	TBD	M	Yes [ ]	Accurate to within −1% unless otherwise stated.

55.12.4 Physical Coding Sublayer (PCS)

Item	Feature	Subclause	Status	Support	Value/Comment
PCT1	The PCS shall	TBD	M	Yes [ ]	

55.12.4.1 PCS receive functions

Item	Feature	Subclause	Status	Support	Value/Comment
PCR1	PCS Receive function shall	TBD	M	Yes [ ]	

55.12.4.2 Other PCS functions

Item	Feature	Subclause	Status	Support	Value/Comment
PCO1	The PCS Reset function shall	TBD	M	Yes [ ]	

55.12.5 Physical Medium Attachment (PMA)

Item	Feature	Subclause	Status	Support	Value/Comment
PMF1	PMA Reset function shall be executed	TBD	M	Yes [ ]	At power on and upon receipt of a reset request from the management entity or from PHY Control.
PMF2	PMA Transmit shall	TBD	M	Yes [ ]	
PMF3	The four transmitters shall be driven by the same transmit clock, TX_TCLK	TBD	M	Yes [ ]	
PMF4	PMA Transmit shall	TBD	M	Yes [ ]	Follow the mathematical description given in TBD.
PMF5	PMA Transmit shall comply with	TBD	M	Yes [ ]	The electrical specifications given in TBD.
PMF6	When the PMA_CONFIG.indicate parameter config is MASTER, the PMA Transmit function shall	TBD	M	Yes [ ]	Source the transmit clock TX_TCLK from a local clock source while meeting the transmit jitter requirements of TBD.
PMF7	When the PMA_CONFIG.indicate parameter config is SLAVE, the PMA Transmit function shall	TBD	M	Yes [ ]	Source the transmit clock TX_TCLK from the recovered clock of TBD while meeting the jitter requirements of TBD.

55.12.6 Management interface

Item	Feature	Subclause	Status	Support	Value/Comment
MF1	All 10GBASE-T PHYs shall provide support for Auto-Negotiation (Clause 28) and shall be capable of operating as MASTER or SLAVE.	TBD	M	Yes [ ]	

55.12.6.1 10GBASE-T Specific Auto-Negotiation Requirements

Item	Feature	Subclause	Status	Support	Value/Comment
AN1	10GBASE-T PHYs shall	TBD	M	Yes [ ]	TBD
AN2	The MASTER-SLAVE relationship shall be determined during Auto-Negotiation	TBD	M	Yes [ ]	TBD

55.12.7 PMA Electrical Specifications

Item	Feature	Subclause	Status	Support	Value/Comment
PME3	The PHY shall provide electrical isolation between	TBD	M	Yes [ ]	The port device circuits including frame ground, and all MDI leads.

55.12.8 Characteristics of the link segment

Item	Feature	Subclause	Status	Support	Value/Comment
LKS1	All implementations of the balanced cabling link shall	TBD	M	Yes [ ]	Be compatible at the MDI.

55.12.9 MDI requirements

Item	Feature	Subclause	Status	Support	Value/Comment
MDI1	MDI connector	TBD	M	Yes [ ]	8-Way connector as per IEC TBD.