UNIVERSITY OF WEST ATTICA

SCHOOL OF ENGINEERING

DEPARTMENT OF INFORMATICS AND COMPUTER ENGINEERING

Master of Science in Science and Technology of Informatics and Computers

MASTER THESIS

Beyond 400G
A Preliminary Study of the 800Gb/s and 1.6Tb/s Technologies

Dimitrios E. Filippou
S.N. 20038

Lecturer: Dr. Antonios Bogris, Professor

MASTER THESIS

Beyond 400G
A Preliminary Study of the 800Gb/s and 1.6Tb/s Technologies

Dimitrios E. Filippou
S.N. 20038

Dimitris Filippou
Digitally signed by Dimitris Filippou
Date: 2023.02.01 16:45:59 +02'00'

Lecturer:

Dr. Antonios Bogris

Antonios Bogris
Digitally signed by Antonios Bogris
Date: 2023.01.31 15:13:55 +02'00'

Examinations Committee:

Panagiotis Karkazis, Associate Professor
Nikolaos Miridakis, Associate Professor

Panagiotis Karkazis
Digitally signed by Panagiotis Karkazis
Date: 2023.02.01 12:15:05 +02'00'

Nikolaos Myridakis
Digitally signed by Nikolaos Myridakis
Date: 2023.01.31 18:32:17 +02'00'

Examinations Date 20/12/2022

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Δημήτριος Φιλίππου του Ευθυμίου, με αριθμό μητρώου MCSE20038 φοιτητής του Προγράμματος Μεταπτυχιακών Σπουδών «Επιστήμη και Τεχνολογία της Πληροφορική και των Υπολογιστών» του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Είμαι συγγραφέας αυτής της μεταπτυχιακής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Επιθυμώ την απαγόρευση πρόσβασης στο πλήρες κείμενο της εργασίας μου εκτός από ακαδημαϊκούς σκοπούς και έπειτα από αίτηση μου στη Βιβλιοθήκη και έγκριση του επιβλέποντα καθηγητή.

Ο Δηλών

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία ολοκληρώθηκε μετά από επίμονες προσπάθειες, σε ένα ενδιαφέρον γνωστικό αντικείμενο, όπως αυτό των σύγχρονων τεχνολογιών μετάδοσης πληροφορίας και ειδικότερα στον κόσμο του Ethernet (IEEE 802.3). Την προσπάθειά μου αυτή υποστήριξε ο επιβλέπων καθηγητής μου, τον οποίο θα ήθελα από καρδίας να ευχαριστήσω.

Ακόμα θα ήθελα να ευχαριστήσω τον εξαίρετο συνάδελφο και Distinguished Engineer της εταιρείας CISCO SYSTEMS, Silicon Valley, κύριο Peter Jones για την πολύτιμη βοήθεια και καθοδήγηση στον κόσμο της τυποποίησης και των πρωτοκόλλων της IEEE.

Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένειά μου, τη σύζυγο και τον γιό μου για την αμέριστη στήριξη τους στην προσπάθεια μου να ολοκληρώσω επιτυχώς τις μεταπτυχιακές μου σπουδές.

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία ασχολείται με την ανάπτυξη αλγορίθμων συγγραφής διπλωματικών εργασιών. Σε αυτό το πεδίο έχουν αναπτυχθεί ιστορικά αρκετοί αλγόριθμοι, κάθε ένας από τους οποίους έχει ορισμένα πλεονεκτήματα και μειονεκτήματα. Ο προτεινόμενος αλγόριθμος λαμβάνει υπ' όψη του τις απαιτήσεις του Τμήματος και αναλαμβάνει να παρουσιάσει διπλωματικές εργασίες οι οποίες να παρουσιάζουν ομοιογένεια και να μην προκαλούν το Γραμματέα ο οποίος θα προσπαθήσει να τις ταξινομήσει στη Βιβλιοθήκη του τμήματος.

Η παρούσα διπλωματική εργασία αποτελεί μια προκαταρκτική μελέτη των τεχνολογιών Ethernet στα 800Gb/s και 1.6Tb/s ως αποτέλεσμα των δραστηριοτήτων της Ομάδας Εργασίας IEEE 802.3 Beyond 400Gb/s Ethernet Task Force (B400G TF), η οποία ιδρύθηκε τον Ιανουάριο του 2021. Οι εργασίες της IEEE B400G TF βρίσκονται σε προκαταρκτικό στάδιο και βασίζονται στην τυπική σειρά προτύπων IEEE Std 802.3. Η IEEE B400G TF ψήφισε τον Νοέμβριο του 2021 το επερχόμενο πρότυπο IEEE P802.3df, το οποίο παρέχει μια αρχιτεκτονική ικανή να υποστηρίξει το Ethernet στα 800Gb/s και 1,6Tb/s, ενώ αναπτύσσει προδιαγραφές Φυσικού Επιπέδου (PHY) για την επικοινωνία σε οπισθεπίπεδα, καλωδιώσεις χαλκού, πολύτροπης και μονότροπης οπτικής ίνας και χρησιμοποιεί την εργασία αυτή για να καθορίσει τις προδιαγραφές των παράγωγων φυσικών επιπέδων και των παραμέτρων διαχείρισης για τη μεταφορά πλαισίων Ethernet στα 100Gb/s, 200Gb/s και 400Gb/s.

Τα τρία πρώτα κεφάλαια παρούσας διπλωματικής εργασίας παρέχουν μια επισκόπηση της αρχιτεκτονικής των IEEE 802.3 προτύπων, σχετικά με τα 100Gb/s, 200Gb/s και 400Gb/s Ethernet. Η αρχιτεκτονική των IEEE 802.3 προτύπων έχει σχέση με το ISO/IEC μοντέλο αναφοράς διασύνδεσης ανοικτών συστημάτων (OSI), ξεκινώντας από το επίπεδο Ελέγχου του Μέσου Πρόσβασης (MAC) έως το PHY, το οποίο αποτελείται από τρία υποστρώματα και έχει πρόσβαση στο μέσο μετάδοσης. Τα κεφάλαια αυτά δίνουν την ευκαιρία να κατανοήσουμε τις απαιτήσεις και τις προδιαγραφές της σειράς των IEEE 802.3 προτύπων, που θα αποτελέσουν τον πρόδρομο ανάπτυξης των επερχόμενων προτύπων IEEE P802.3df και IEEE P802.3dj.

Το τέταρτο κεφάλαιο παρέχει τις αρχιτεκτονικές προοπτικές για την ανάπτυξη των IEEE P802.3df και IEEE P802.3dj, μέσω των προτάσεων που παραδόθηκαν από τα

μέλη της IEEE B400G TF. Οι προτάσεις αυτές αποτελούν μέρος της έρευνας της παρούσας διπλωματικής εργασίας και των τελευταίων εξελίξεων στην τεχνολογία του Ethernet. Η δομή του κεφαλαίου αυτού ακολουθεί τη μορφή της τυπικής σειράς των IEEE 802.3 προτύπων, παρέχει ανάλυση βάσει του εκάστοτε επιπέδου και δίνει την ευκαιρία για μια ολιστική άποψη των επερχόμενων προοπτικών.

Τα παραρτήματα Α έως Μ παρατίθενται στην παρούσα διπλωματική εργασία για ενημερωτικούς σκοπούς, καθώς αποτελούν μέρος των IEEE 802.3 προτύπων.

ABSTRACT

This paper provides a preliminary study of 800Gb/s and 1.6Tb/s Ethernet technologies as a result of the activities of the IEEE 802.3 Beyond 400Gb/s Ethernet Task Force (B400G TF), which was established in January 2021. The works of IEEE B400G TF are in preliminary level and are based on the IEEE Std 802.3 standard series. IEEE B400G TF voted in November 2021 the upcoming IEEE P802.3df standard, which provides an architecture capable of supporting both 800Gb/s and 1.6Tb/s Ethernet, while producing Physical Layer (PHY) specifications for communication across backplanes, copper cabling, multimode and single-mode fiber and use this work to define derivative physical layer specifications and management parameters for the transfer of Ethernet format frames at 100Gb/s, 200Gb/s and 400Gb/s.

The first three chapters of this study provide an overview of the IEEE 802.3 architecture, regarding 100Gb/s, 200Gb/s and 400Gb/s Ethernet. IEEE 802.3 architecture has a relationship to the ISO/IEC Open System Interconnection (OSI) reference model, starting from the Media Access Control (MAC) layer to the PHY, which consists of three sublayers and has access to the medium. These chapters give the opportunity to understand the requirements and specifications of IEEE 802.3 standard series, which will be the development precursor of the upcoming IEEE P802.3df and IEEE P802.3dj standards.

The fourth chapter provides the architectural perspectives for the development of IEEE P802.3df and IEEE P802.3dj, through the proposals delivered from the IEEE B400G TF members. These proposals are part of the research and the latest developments in Ethernet technology. The structure of this chapter follows the format

of the IEEE 802.3 standard series, provides a layer-based analysis, and gives the opportunity for a holistic view of the upcoming perspectives.

Annexes A to M are included in this paper for informative purposes, since they are part of the IEEE 802.3 standards.

SCIENTIFIC DOMAIN: IEEE Standards for Ethernet, Media Access Control Parameters, Physical Layers, and Management Parameters

KEYWORDS: 100Gb/s Ethernet, 100CGMII, 100GAUI, 100GBASE-ER4, 100GBASE-R, 100GBASE-LR4, 100GBASE-SR10, 200Gb/s Ethernet, 200GAUI-4, 200GAUI-8, 200GBASE-DR4, 200GBASE-FR4, 200GBASE-LR4, 200GBASE-R, 200GMII, 200GXS, 400Gb/s Ethernet, 400GAUI-8, 400GAUI-16, 400GBASE-DR4, 400GBASE-FR8, 400GBASE-LR8, 400GBASE-SR16, 400GBASE-R, 400GMII, 400GXS, 800Gb/s Ethernet, 1.6Tb/s Ethernet, EEE, Energy Efficient Ethernet, Ethernet, FEC, Forward Error Correction, IEEE 802.3, IEEE 802.3bs, IEEE 802.3df, IEEE 802.3dj, MDI, Medium Dependent Interface, MMF, MultiMode Fiber, PAM4, PCS, Physical Coding Sublayer, PMA, Physical Medium Dependent, PMD, Physical Medium Dependent sublayer, RS, Reconciliation Sublayer, SMF, Single Mode Fiber.

Contents

FIGURES

TABLES

Abbreviations

ADC	Analogue-to-Digital Converter
AN	Auto-Negotiation
ANR	Amplitude-to-Noise Ratio
ASIC	Application-Specific Integrated Circuit
BCH	Bose, Ray-Chaudhuri, and Hocquenghem
BCJR	Bahl, Cocke, Jelinek, Raviv
BER	Bit Error Ratio
BPS	Blind Phase Search
BRE	Bridge Relay Entity,
BT	Bit Time
C2C	Chip-to-Chip
C2M	Chip-to-Module
CAUI-4	100Gb/s four-lane Attachment Unit Interface
CAUI-10	100Gb/s ten-lane Attachment Unit Interface
CD	Chromatic Dispersion
CFC	Carrier Frequency Compensation
CFEC	Concatenated FEC
COM	Channel Operating Margin
CPC	Carrier Phase Compensation
CPPI	100Gb/s Parallel Physical Interface
CRU	Clock Recovery Unit
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
CTLE	Continuous Time Linear Equalizer
DD	Direct Detect
DDJ	Data Depend Jitter
DDPWS	Data Dependent Pulse Width Shrinkage
DFE	Decision Feedback Equalizer
DGD	Differential Group Delay
DIC	Deficit Idle Counter
DLF	Delay Line Filter
DLL	Data Link Layer
DSP	Digital Signal Processor

DTE	Data Terminal Equipment
DUT	Device Under Test
DWDM	Dense Wavelength Division Multiplexing
EEE	Energy Efficient Ethernet
EQ	EQualization
FCS	Frame Check Sequence
FEC	Forward Error Correction
FFE	Feed-Forward Equalizer
CGMII	100Gb/s Media Independent Interface
HCB	Host Compliance Board
HD-FEC	Hard-Decision FEC
IEC	International Electrotechnical Committee
IL	Insertion Loss
ILD	Insertion Loss Deviation
IPG	InterPacket Gap
ISI	InterSymbol Interference
ISO	International Standardization Organization
JTol	Jitter Tolerance
LAN	Local Area Network
LB	Loop Bandwidth
LLC	Logical Link Control
LMS	Least-Mean Square
LPI	Low Power Idle
MAC	Media Access Control
MAN	Metropolitan Area Network
MCB	Module Compliance Board
MDC	Management Data Clock
MDI	Medium Dependent Interface
MDIO	Management Data Input/Output
MIB	Management Information Base
MII	Media Independent Interface
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood

MLSD	Maximum Likelihood Sequence Detector
MLSE	Maximum Likelihood Sequence Estimation
MMD	Management Data Input/Output (MDIO) Manageable Device
MMF	MultiMode Fiber
MTTFPA	Mean Time To False Packet Acceptance
MUX	Multiplexer
NCG	Net Coding Gain
nPPI	n-Parallel Physical Interface
NRZ	Non-Return-to-Zero
OFEC	Open FEC
OMA	Optical Modulation Amplitude
OSI	Open System Interconnection
PAM4	4-level Pulse Amplitude Modulation
PCS	Physical Coding Sublayer
PCSL	PCS Lane
PD	PhotoDetector
PFC	Priority-based Flow Control
PHY	Physical Layer
PI	Phase Interpolator
PLL	Phase Locked Loop
PLS	Physical Signaling Sublayer
PMA	Physical Medium Attachment
PPI	Parallel Physical Interface
PR-PAM4	Partial Response PAM4
PSD	Power Spectrum Density
QSFP-DD	Quad Small Form-Factor – Double Density
RAM	Rapid Alignment Marker
RIN	Relative Intensity Noise
RIN_xOMA	Relative Intensity Noise Optical Modulation Amplitude
RS	Reconciliation Sublayer
SD	Soft-Decision
SE	Single-Ended
SE-PAM4	Single-Ended PAM4

SECQ	Stressed Eye Closure for PAM4
SER	Symbol Error Ration
SerDes	Serializer/Deserializer
SFD	Start Frame Delimiter
SD-FEC	Soft-Decision FEC
SMF	Single-Mode Fiber
SMSR	Side-Mode Suppression Ratio
SNR	Signal to Noise Ratio
SP	Skew Point
SOVA	Soft-decision Viterbi Algorith
SPC	Single Parity Check
TDECQ	Transmitter and Dispersion Eye Closure for PAM4
TDP	Transmitter and Dispersion Penalty
TIA	TransImpedance Amplifier
TP	Test Point
VECP	Vertical Eye Closure Penalty
VCO	Voltage-Controlled Oscillator
VMA	Voltage Modulation Amplitude
XS	eXtender Sublayer

1 Introduction

IEEE Std 802.3 is an international standard series for Local and Metropolitan Area Networks (LANs and MANs), employing Carrier Sense Multiple Access with Collision Detection (CSMA/CD) as the shared media access method, the IEEE 802.3 protocol and frame format for data communication. These international standard series is intended to encompass several media types and techniques for a variety of MAC data rates. (IEEE Std 802.3™, 2018)

IEEE 802.3 standard series defines Ethernet local area, access, and metropolitan area networks. Ethernet is specified at selected speeds of operation and uses a common MAC specification and Management Information Base (MIB). The CSMA/CD MAC protocol specifies a shared medium half duplex operation, as well as full duplex operation. Speed specific Media Independent Interfaces (MIIs) provide an architectural and optional implementation interface to selected PHY entities. The PHY encodes the frames for transmission and decodes the received frames with the modulation specified for the speed of operation, the transmission medium and the supported link length. Other specified capabilities include the control and management protocols, and the provision of power over selected twisted pair PHY types.

In 2017, IEEE published IEEE Std 802.3bs-2017 (IEEE STANDARDS ASSOCATION, 2017) which describes the general requirements for 200Gb/s and 400Gb/s Ethernet. 200 Gigabit Ethernet uses the IEEE 802.3 MAC sublayer operating at a data rate of 200Gb/s, coupled with any IEEE 802.3 200GBASE PHY implementation, and is defined for full duplex operation only. 400 Gigabit Ethernet uses the IEEE 802.3 MAC sublayer operating at a data rate of 400 Gb/s, coupled with any IEEE 802.3 400GBASE PHY implementation, and is defined also for full duplex operation only.

In 2018, IEEE published IEEE Std 802.3™-2018 (IEEE Std 802.3™, 2018), which describes the general requirements for starting from 10Gb/s to 400Gb/s Ethernet. For example, IEEE Std 802.3u™ added 100 Mb/s operation (also called Fast Ethernet), IEEE Std 802.3z added 1000 Mb/s operation (also called Gigabit Ethernet), IEEE Std 802.3ae added 10Gb/s operation (also called 10 Gigabit Ethernet), IEEE Std 802.3ah™ specified access network Ethernet (also called Ethernet in the First Mile) and IEEE Std 802.3ba added 40Gb/s operation (also called 40 Gigabit Ethernet) and

100Gb/s operation (also called 100 Gigabit Ethernet). IEEE Std 802.3™-2018 also includes the updated specifications of the IEEE Std 802.3bs-2017.

Note: The information provided in this study is only for academic use and not for any other purposes.

2 IEEE 802.3™ - 2018 Ethernet Standard

2.1 Media Access Control (MAC) Sublayer

The MAC sublayer defines a medium-independent capacity, based on the medium-dependent physical capacity, which provided by the PHY, and it is under the access-layer-independent LAN LLC sublayer (or another MAC client). This is applicable to a range of local area broadcast media suitable for the media access discipline known as CSMA/CD.

The LLC and the MAC sublayers are intended to have the same functionality as that described in the OSI model for the Data Link Layer (DLL). In a broadcast network, the concept of a data link between two network entities does not correspond directly to a distinct physical connection. Nevertheless, the partitioning of functions requires two main functions generally associated with a data link control procedure to be performed in the MAC sublayer. These functions are the following:

1. Data encapsulation (transmit and receive)
    a. Framing (frame boundary delimitation, frame synchronization)
    b. Addressing (handling of source and destination addresses)
    c. Error detection (detection of physical medium transmission errors)
2. Media Access Management
    a. Medium allocation (collision avoidance)
    b. Contention resolution (collision handling)

The optional MAC control sublayer, architecturally positioned between LLC (or other MAC client) and the MAC is transparent to both the underlying MAC and its client (typically LLC). The MAC sublayer operates independently of its client, e.g., it is unaware whether the client is LLC or the MAC Control sublayer. This allows the MAC to be specified and implemented, either the MAC Control sublayer is implemented or not.

IEEE 802.3™ provides for two modes of operation of the MAC sublayer:

1. The half-duplex mode, where stations contend for the use of the physical medium, using the CSMA/CD algorithms specified. Bidirectional communication is implemented by the fast exchange of frames, rather than the full duplex operation. The half-duplex operation can be used with all supported media, and it is required with those media that are unable to support simultaneous transmission and reception without interference.

2. The full duplex mode of operation can be used when all the following are true:
    a. The physical medium can support simultaneous transmission and reception without interference.
    b. The physical medium between two stations on a LAN can be treated as a full duplex point-to-point link. So, when there is no contention for use of the shared medium, the multiple access (e.g., CSMA/CD) algorithms are not necessary.
    c. Both stations on the LAN are capable of and have been configured to use full duplex operation.

The most common configuration envisioned for full duplex operation consists of a switch with a dedicated LAN connecting each port to a single device.

MAC sublayer provides for real-time control and manipulation of MAC sublayer operation. MAC sublayer specifies a generalized architecture and protocol for MAC Control. The MAC Control protocol is specified such that it can support new functions to be implemented and added to IEEE 802.3™ standard in the future.

Non-real-time or quasistatic control (e.g., configuration of MAC operational parameters) is provided by the layer management. The operation of the MAC Control sublayer is transparent to the CSMA/CD MAC.

2.1.1 MAC Control Sublayer Architecture

The MAC control sublayer is client of the CSMA/CD MAC. ***Figure 2-1*** depicts the architectural positioning of the MAC Control sublayer with respect to the CSMA/CD MAC and the MAC Control client. MAC Control clients may include the Bridge Relay Entity (BRE), LLC) or other applications.

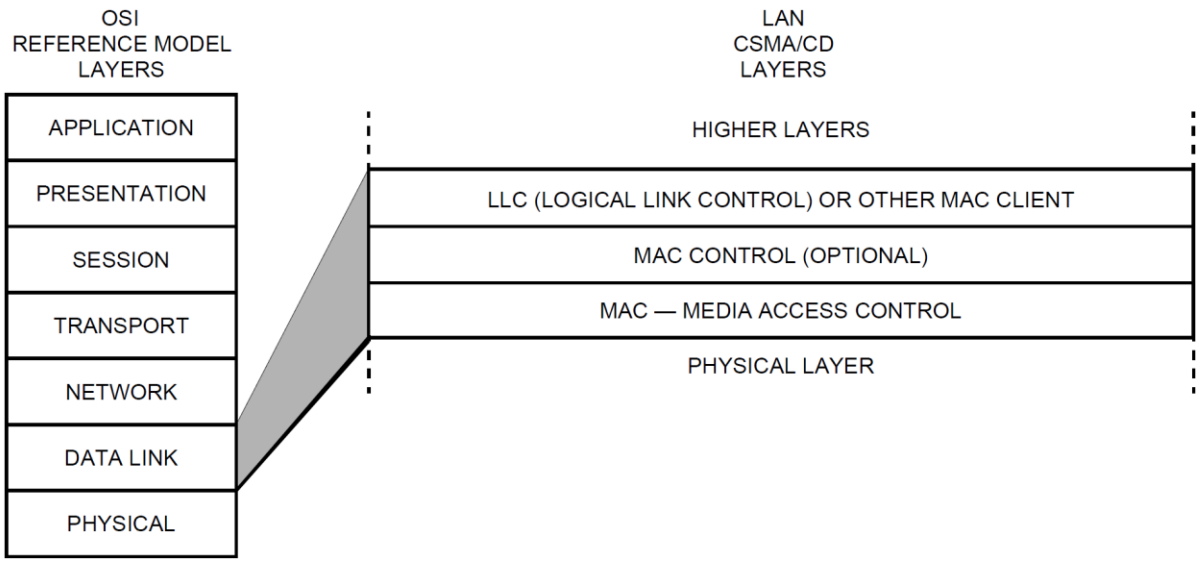


Figure 2-1. Architectural positioning of MAC Control sublayer.

The optional MAC Control sublayer is allocated between the MAC sublayer and its MAC client. The MAC Control sublayer uses the MAC service interface to interface to the MAC client and to the MAC. **Figure 2-2** depicts the usage of interlayer interfaces by the MAC Control sublayer. Devices that implement the MAC Control sublayer shall support the MAC service primitives, *MA_CONTROL.request* and *MA_-CONTROL.indication*.

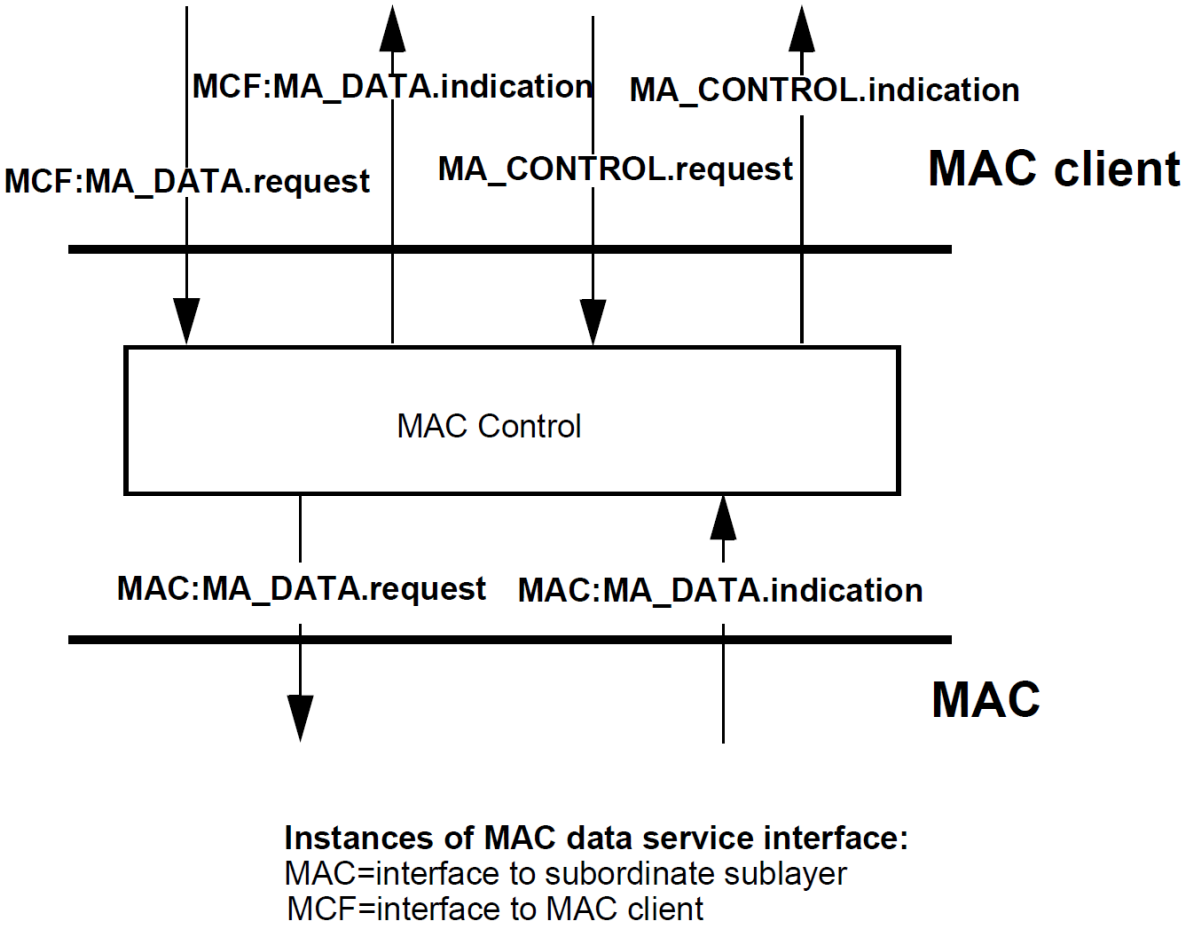


Figure 2-2. *MAC Control sublayer supports of interlayer service interfaces.*

The services provided by the MAC sublayer allow the local MAC client entity to exchange LLC data with peer LLC sublayer entities.

2.1.2 MAC Control Frame Format

MAC Control frames are fixed length, containing *minFrameSize*–32 bits. The underlying MAC prepends the Preamble and Start-of-Frame delimiter fields and appends the FCS. **Figure 2-3** depicts the MAC Control frame format.

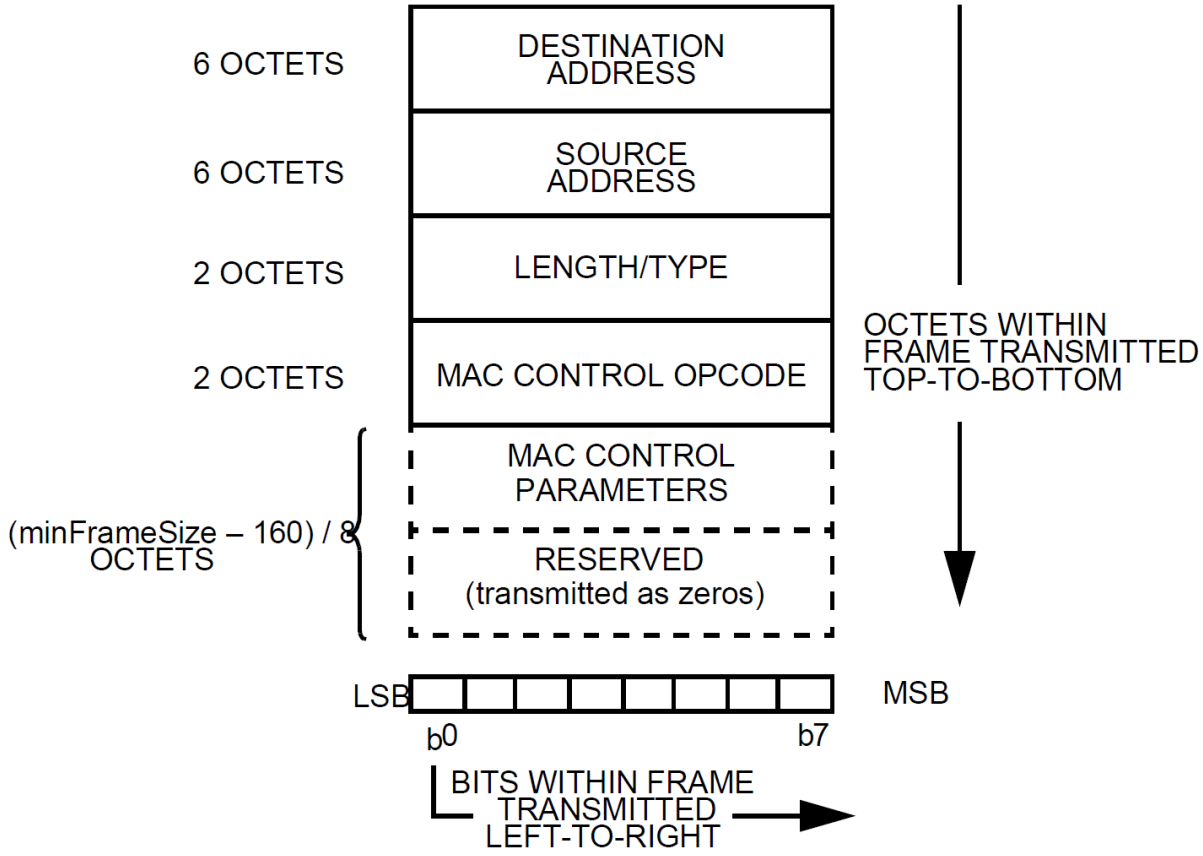


Figure 2-3. *MAC Control frame format.*

The Destination Address field of the MAC Control frame contains the 48-bit address of the station(s) for which the frame is intended. It can be either an individual or a multicast (including broadcast) address. The permitted values of the Destination Address field can be specified separately for each MAC Control opcode.

The Source Address field of the MAC Control frame contains the 48-bit individual address of the station sending the frame.

The Length/Type field of the MAC Control frame is a 2-octet field that contains the hexadecimal value 88-08. This value carries the Ethertype interpretation and has been universally assigned for MAC Control of CSMA/CD LANs.

The MAC Control Opcode field contains a 2-octet operation code, which indicates the MAC Control function. It defines the semantics of the MAC Control Parameters field, which contains the list of defined MAC Control opcodes and interpretations. The MAC Control frame contains exactly one MAC Control opcode.

The MAC Control Parameters field contains the MAC Control opcode-specific parameters. This field may contain none, one, or more parameters as defined by the MAC Control Opcode. The opcode-specific semantics of the MAC Control Parameters

field specify each MAC Control function. The MAC Control Parameters field contains an integral number of octets. The length of the MAC Control Parameters field varies from a minimum of zero, to a maximum of minFrameSize - 160 bits.

The Reserved field is used only when the MAC Control parameters do not fill the fixed length MAC Control frame. The size of the Reserved field, which is optional, is determined by the size of the MAC Control Parameters field provided by the MAC Control and the minimum frame size parameter of the particular implementation. The length of Reserved field required for a MAC Control Parameters field that is n octets long is [minFrameSize – (8 × n + 160)] bits. The Reserved field is transmitted as all zeros.

2.2 Architectural Perspectives of 100Gb/s Ethernet

100 Gigabit Ethernet uses the IEEE 802.3 MAC sublayer, which operates at a data rate of 100Gb/s, coupled with any IEEE 802.3 100GBASE PHY implementation. 100Gb/s PHY entities, such as those specified in **Table 2-1**, provide a Bit Error Ratio (BER) better than or equal to 10^{-12} at the MAC/ Physical Signaling Sublayer (PLS) service interface. 100 Gigabit Ethernet is designed for a full duplex operation only.

100 Gigabit Ethernet couples the IEEE 802.3 MAC to a family of 100Gb/s PHYs. The relationship among 100 Gigabit Ethernet, the IEEE 802.3 MAC, and the ISO/IEC OSI reference model are shown in **Figure 2-4**.

While this specification defines interfaces in terms of bits, octets, and frames, implementations can select another data-path width for implementation convenience. The only exceptions are as follows:

a. The CGMII (100Gb/s Media Independent interface), which, when implemented as a logical interconnection port between the MAC sublayer and the PHY, uses a 64-bit wide data path. Physical instantiations of this interface can use another data-path width.
b. The management interface, which, when physically implemented as the Management Data Input/Output and Management Data Clock (MDIO/MDC) at an observable interconnection port, uses a bit-wide data path.
c. The Physical Medium Attachment (PMA) service interface, which, when physically implemented as CAUI-4 (100Gb/s four-lane Attachment Unit Interface) at an observable interconnection port, uses a 4-lane data path, as specified in **Annex H**, **Annex I**, **Annex K**, or **Annex L**.

d. The PMA service interface, which, when physically implemented as CAUI-10 (100Gb/s ten-lane Attachment Unit Interface) at an observable interconnection port, uses a 10-lane data path as specified in **Annex H** or **Annex I**.
e. The PMD service interface, which, when physically implemented as CPPI (100Gb/s Parallel Physical Interface) at an observable interconnection port, uses a 10-lane data path as specified in **Annex M**.
f. The Medium Dependent Interfaces (MDIs) as specified in Section **2.7** for 100GBASE-LR4 and 100GBASE-ER4, and in Section **2.9** for 100GBASE-SR4 all use a 4-lane data path.
g. The MDIs as specified in Section **2.6** for 100GBASE-SR10 use a 10-lane data path.

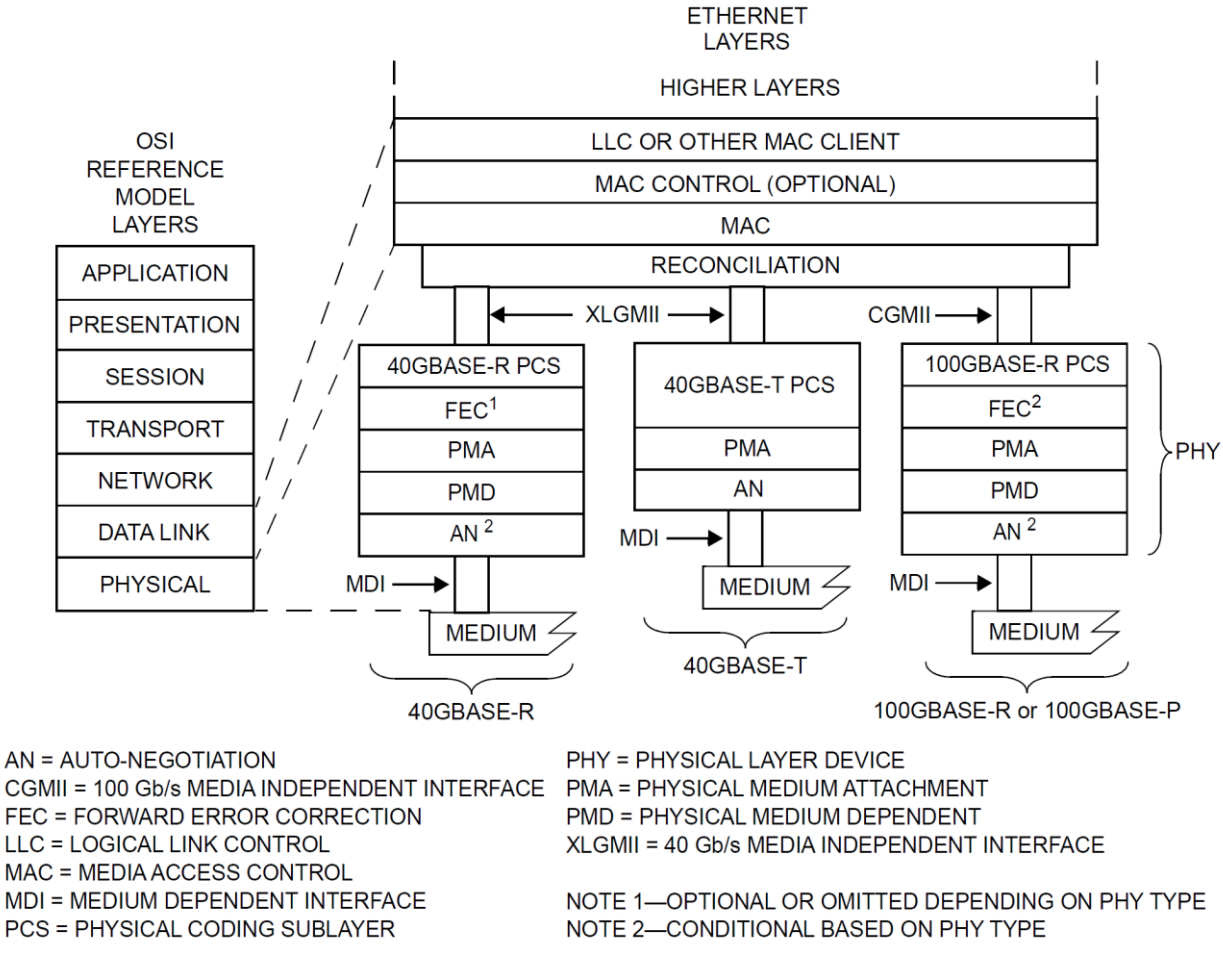


Figure 2-4. *Architectural positioning of 100Gb/s Ethernet.*

The nomenclature employed by the 100Gb/s PHYs is explained as follows.

The alpha-numeric prefix 100GBASE in the port type (e.g., 100GBASE-R) represents the family of PHY devices operating at the speed of 100Gb/s.

100GBASE-R represents a family of PHY devices using the Section **2.4** Physical Coding Sublayer (PCS) for 100Gb/s operation over multiple PCS lanes (see Section **2.4**) and a PMD implementing 2-level Pulse Amplitude Modulation (PAM). Some 100GBASE-R PHY devices also use the transcoding and Forward Error Correction (FEC) of Section **2.8**.

100GBASE-R represents PHY devices using the Section **2.4** PCS for 100Gb/s operation over multiple PCS lanes (see Section **2.4**) and a PMD implementing more than 2-level PAM. Some 100GBASE-P PHY devices also use the transcoding and FEC of Section **2.8**.

PHY devices listed in **Table 2-1** are defined for operation at 100Gb/s.

Name	Description
100GBASE-SR10	100Gb/s PHY using 100GBASE-R encoding over ten lanes of multimode fiber, with reach up to at least 100m (see Section **2.6**)
100GBASE-SR4	100Gb/s PHY using 100GBASE-R encoding over four lanes of multimode fiber, with reach up to at least 100m (see Section **2.9**)
100GBASE-LR4	100Gb/s PHY using 100GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 10km (see Section **2.7**)
100GBASE-ER4	100Gb/s PHY using 100GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 40km (see Section **2.7**)

Table 2-1. *100Gb/s PHYs.*

2.2.1 Delay constraints

The predictable operation of the MAC Control *PAUSE* operation demands an upper bound on the propagation delays through the network. This implies that the MAC, the MAC Control sublayer, and the PHY implementers must conform to certain delay maxima, and the network planners and administrators conform to constraints regarding the cable topology and the concatenation of devices. **Table 2-2** depicts the values of the maximum sublayer delay (the sum of transmit and receive delays at one end of the link) in bit times, where *pause_quanta* is the unit of measurement for the pause time specified as 512 MAC bit times as specified in **Annex G**. If the PHY

contains an Auto-Negotiation sublayer, the delay of this sublayer is included within the delay of the PMD and the medium.

Sublayer	Maximum (bit time)	Maximum (pause_quanta)	Maximum (ns)	Notes
100G MAC, RS, and MAC Control	24 576	48	245.76	
100GBASE-R PCS	35 328	69	353.28	
100GBASE-R FEC	122 880	240	1228.8	
100GBASE-R RS-FEC	40 960	80	409.60	
100GBASE-R PMA	9 216	18	92.16	
100GBASE-SR4 PMD	2 048	4	20.48	Includes 2 m of fiber.
100GBASE-SR10 PMD	2 048	4	20.48	Includes 2 m of fiber.
100GBASE-LR4 PMD	2 048	4	20.48	Includes 2 m of fiber.
100GBASE-ER4 PMD	2 048	4	20.48	Includes 2 m of fiber.

Table 2-2. *Sublayer delay constraints.*

Equation 2-1 specifies the calculation of cable delay in nanoseconds per meter of fiber cable, based upon the parameter *n*, which represents the ratio of the speed of electromagnetic propagation in the fiber cable to the speed of light in a vacuum, c = 3 × 10^8 m/s.

$$cable\ delay = \frac{10^9}{n \cdot c} ns/m$$

Equation 2-1

The value of *n* should be available from the fiber cable manufacturer, but if no value is known, then a conservative delay estimate can be calculated using a default value of *n*=0.66, which yields a default cable delay of 5ns/m.

2.2.2 Skew constraints

Skew (or relative delay) can be introduced between the lanes by both the active and the passive elements of a 100GBASE-R link. Skew is defined as the difference between the times of the earliest and the latest PCS lanes for the one to zero transition of the alignment marker sync bits. The PCS deskew function compensates for all lane-to-lane the Skew observed at the receiver. The Skew between the lanes must be kept within the limits of the **Table 2-3**, so that the transmitted information on the lanes can be reassembled by the receive PCS.

Skew points	Maximum Skew (ns) [a]	Maximum Skew for 100GBASE-R PCS lane (UI) [b]
SP0	29	≈150
SP1	29	≈150
SP2	43	≈222
SP3	54	≈278
SP4	134	≈691
SP5	145	≈748
SP6	160	≈824
SP7	29	≈150
At PCS receive	180	≈928
At RS-FEC transmit	49	≈253
At RS-FEC receive [c]	180	≈4641
At PCS receive (with RS-FEC)	49	≈253

[a] The Skew limit includes 1ns allowance for PCB traces that are associated with the Skew points.

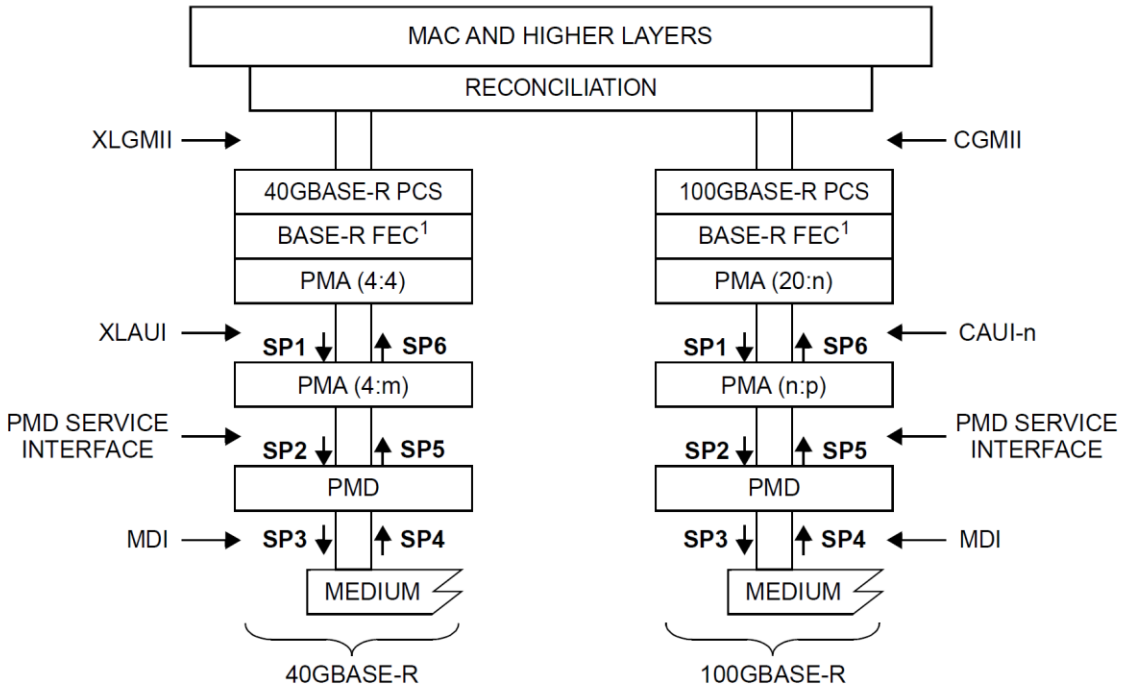
[b] The symbol ≈ indicates the approximate equivalent of maximum Skew in UI for 100GBASE-R, where 1 UI equals 193.939394ps at PCS lane signaling rate of 5.15625GBd.

[c] The skew at the RS-FEC receive is the skew between RS-FEC lanes. The symbol ≈ indicates approximate equivalent of maximum Skew in UI for RS-FEC lanes with a signaling rate of 25.78125GBd.

Table 2-3. *Summary of Skew constraints.*

Skew Variation can be introduced due to variations in electrical, thermal or environmental characteristics. Skew Variation is defined as the change in Skew between any PCS lane and any other PCS lane over the entire time that the link is in operation. From the time the link is brought up, the Skew Variation must be limited to ensure that each PCS lane always traverses the same lane between any pair of adjacent sublayers, while the link remains in operation.

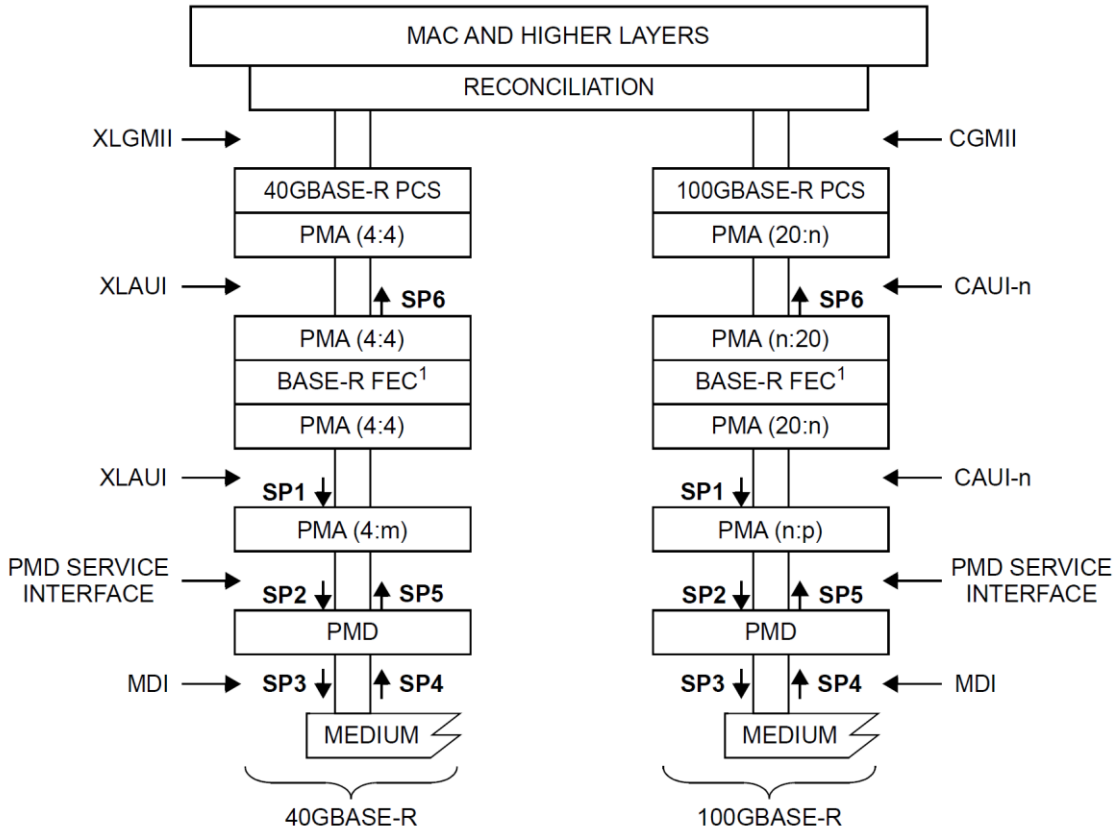
The maximum Skew and the Skew Variation at physically instantiated interfaces are specified at the Skew points SP1, SP2, and SP3 for the transmit direction and at the Skew points SP4, SP5, and SP6 for the receive direction as shown in **Figure 2-5** (single CAUI-n interface) and in **Figure 2-6** (multiple CAUI-n interfaces).



CAUI-n = 100 Gb/s ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLAUI = 40 Gb/s ATTACHMENT UNIT INTERFACE
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

m = 1 or 4
n = 4 or 10
p = 4 or 10
NOTE1—CONDITIONAL BASED ON PHY TYPE

Figure 2-5. *100GBASE-R Skew points for single CAUI-n (40BASE-R is out of the scope of the study).*



CAUI-n = 100 Gb/s ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLAUI = 40 Gb/s ATTACHMENT UNIT INTERFACE
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

m = 1 or 4
n = 4 or 10
p = 4 or 10
NOTE1—CONDITIONAL BASED ON PHY TYPE

Figure 2-6. *100GBASE-R Skew points for multiple CAUI-n (40BASE-R is out of the scope of the study).*

In the transmit direction, the Skew points are defined on the following locations:

- SP1 on the CAUI-n interface, at the input of the PMA closest to the PMD,
- SP2 on the PMD service interface, at the input of the PMD,
- SP3 at the output of the PMD, at the MDI.

In the receive direction, the Skew points are defined on the following locations:

- SP4 at the MDI, at the input of the PMD,
- SP5 on the PMD service interface, at the output of the PMD,
- SP6 on the CAUI-n interface, at the output of the PMA closest to the PCS.

The allowable limits for Skew are shown in **Table 2-3** and the allowable limits for Skew Variation are shown in **Table 2-4**.

Skew points	Maximum Skew Variation (ns)	Maximum Skew Variation for 25.78125GBd PMD lane (UI) [a]
SP0	0.2	N/A
SP1	0.2	N/A
SP2	0.4	≈10
SP3	0.6	≈15
SP4	3.4	≈88
SP5	3.6	≈93
SP6	3.8	≈98
SP7	0.2	N/A
At PCS receive	4	N/A
At RS-FEC transmit	0.4	≈10
At RS-FEC Receive [b]	4	≈103
At PCS receive (with RS-FEC)	0.4	≈10

[a] The symbol ≈ indicates the approximate equivalent of maximum Skew Variation in UI for 100GBASE-R, where 1 UI equals 38.787879ps at PMD lane signaling rate of 25.78125GBd.

[b] The skew at the RS-FEC receive is the skew between RS-FEC lanes.

Table 2-4. *Summary of Skew Variation constraints.*

The skew points are illustrated for a PHY incorporating RS-FEC in **Figure 2-7**.

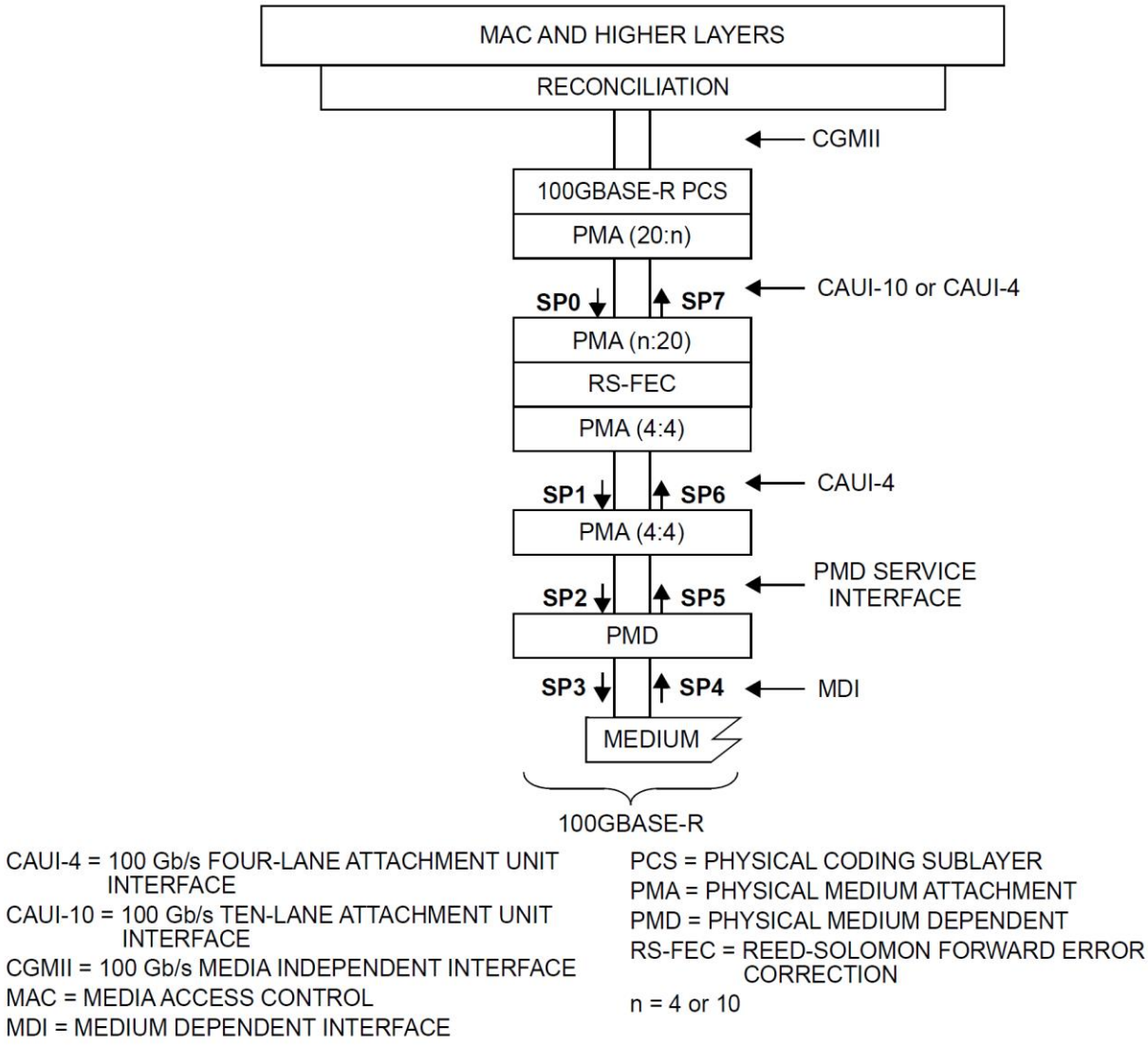


***Figure 2-7**. 100GBASE-R Skew points with RS-FEC and CAUI-n.*

The Skew requirements for the PCS, PMA and the PMD sublayers are specified in **Table 3-5** and **Table 3-6**.

2.3 Reconciliation Sublayer (RS) and Media Independent Interface for 100Gb/s operation (CGMII)

The Reconciliation Sublayer (RS) and the MII are between the Ethernet media access controllers and the various PHYs. **Figure 2-8** shows the relationship of the RS and MII to the ISO/IEC OSI reference model.

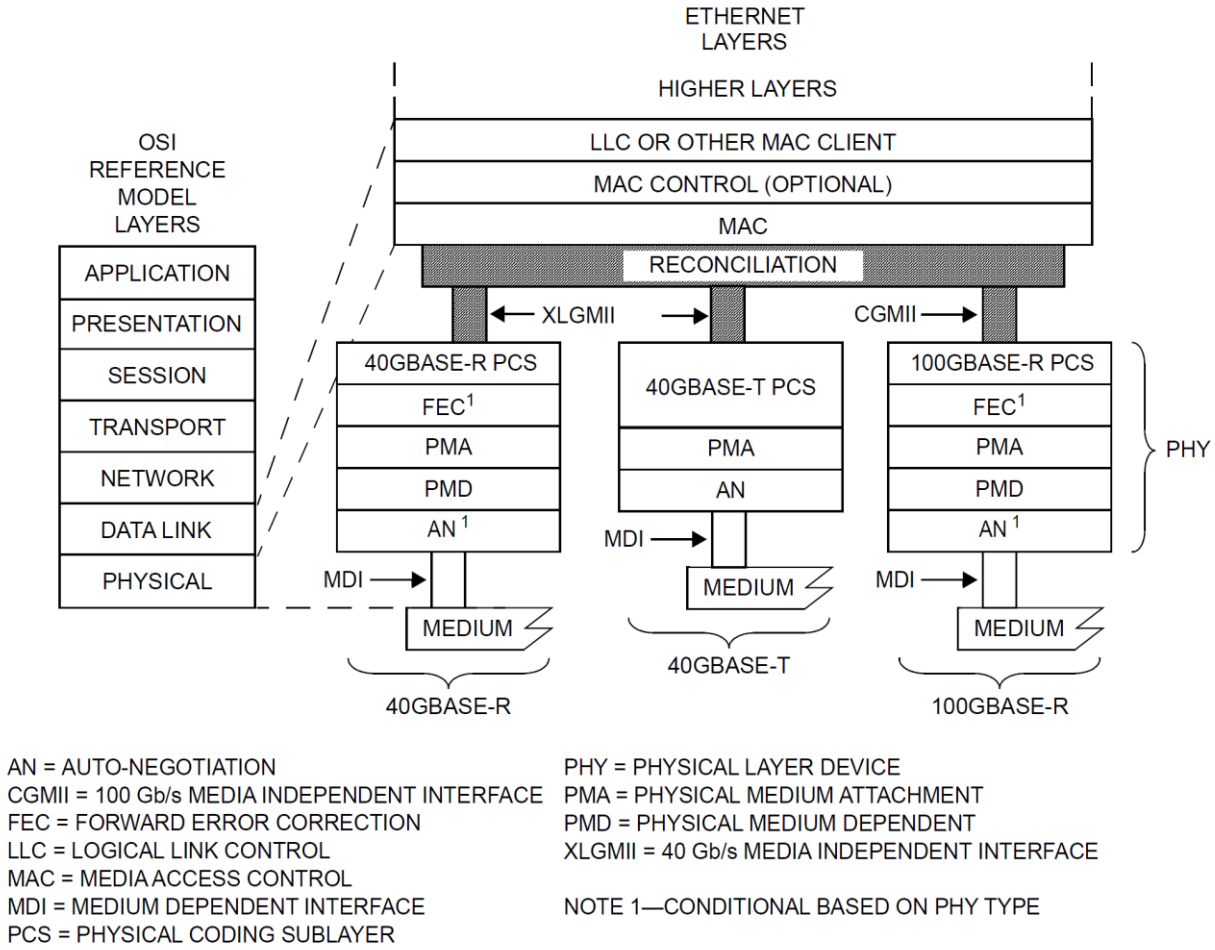


Figure 2-8. *RS and MII relationship to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model.*

The CGMII are optional logical interfaces between the MAC sublayer and the PHY.

The RS adapts the bit serial protocols of the MAC to the parallel format of the PCS service interface. The PCS is specified to the CGMII, so if not implemented, a conforming implementation shall operate functionally as if the RS and CGMII were implemented.

The CGMII has the following characteristics:

a. The CGMII supports a speed of 100 Gb/s.
b. Data and delimiters are synchronous to a clock reference.
c. It provides independent 64-bit wide transmit and receive data paths.
d. It supports full duplex operation only.

The major concepts of the CGMII are the following:

a. The CGMII is functionally similar to other media independent interfaces that have been defined for lower speeds, as they all define an interface allowing independent development of MAC and PHY logic.

b. The RS converts between the MAC serial data stream and the parallel data paths of the CGMII.
c. The RS maps the signal set provided at the CGMII to the PLS service primitives provided at the MAC.
d. Each direction of data transfer is independent and serviced by data, control, and clock signals.
e. The RS generates continuous data or control characters on the transmit path and expects continuous data or control characters on the receive path.
f. The RS participates in link fault detection and reports by monitoring the receive path for status reports that indicate an unreliable link and generates status reports on the transmit path in order to report detected link faults to the DTE on the remote end of the connecting link.
g. The CGMII may also support Low Power Idle (LPI) signaling for PHY types supporting Energy Efficient Ethernet (EEE) (see **Annex F**).

The CGMII is specified to support 100Gb/s operation.

2.3.1 Delay constraints

The maximum cumulative MAC Control, MAC, and RS delay (sum of transmit and receive delays at one end of the link) shall meet the values specified in **Table 2-5**. A description of overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **2.2.1** and its references.

Sublayer	Maximum (bit time)	Maximum (pause_quanta)	Maximum (ns)
100 Gb/s MAC, RS, and MAC Control	24576	48	245.76

Table 2-5. *Delay constraints.*

The functions allocated at the CGMII balance the need for media independence with interface simplicity. The CGMII maximize media independence by separating the Data Link and PHYs of the OSI seven-layer reference model.

The CGMII is composed of independent transmit and receive paths. Each direction uses 64 data signals (*TXD<63:0> and RXD<63:0>*), 8 control signals (*TXC<7:0> and RXC<7:0>*), and a clock (*TX_CLK and RX_CLK*). **Figure 2-9** depicts a schematic view of the RS inputs and outputs.

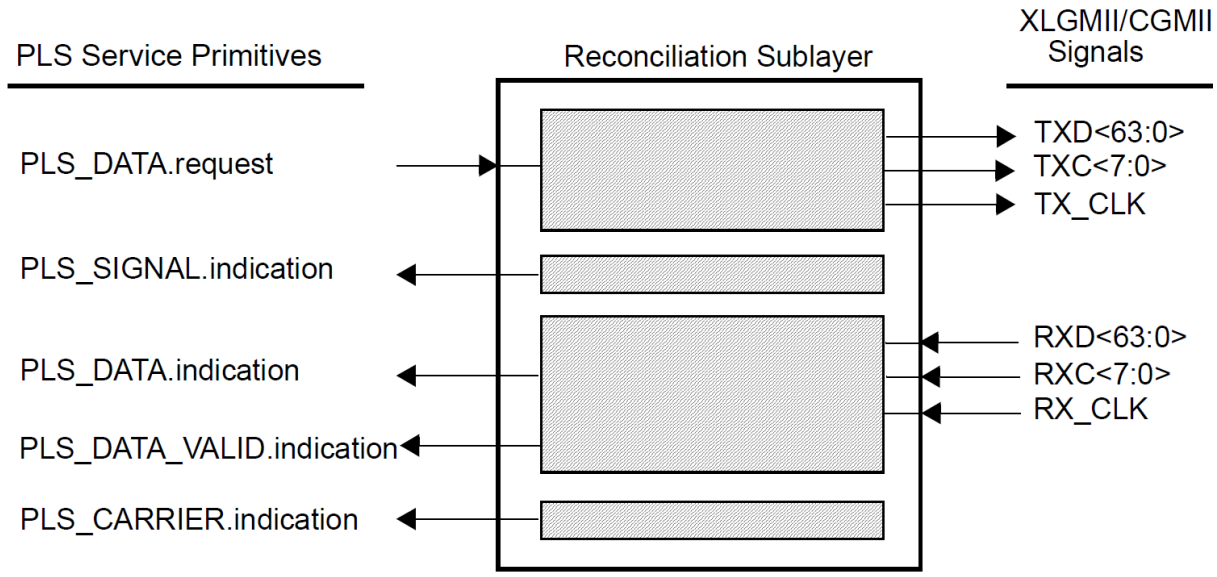


Figure 2-9. *Reconciliation Sublayer (RS) inputs and outputs.*

2.4 Physical Coding Sublayer (PCS) for 64B/66B, type 100GBASE-R

The PCS is common to two families of 100Gb/s PHY implementations, known as 100GBASE-R. The 100GBASE-R PCS is a sublayer of the 100Gb/s PHYs listed in **Table 2-1**. The term 100GBASE-R is used when referring generally to PHYs using the PCS defined in this Section.

The 100GBASE-R is based on a 64B/66B code. The 64B/66B code supports the data and control characters transmission, while maintaining robust error detection. Data distribution is introduced to support multiple lanes in the PHY. Part of the distribution includes the periodic insertion of an alignment marker, which allows the receive PCS to align data from multiple lanes.

Figure 2-10 depicts the 100GBASE-R sublayer (shown shaded), the Ethernet MAC and Reconciliation Sublayers, and the higher layers (40GBASE-R is out of the scope of this study).

64B/66B encoding is reused with appropriate changes made to support 8-byte alignment versus 4-byte alignment. In addition to 64B/66B encoding, there is a methodology to add alignment markers and distribute data to multiple lanes.

Figure 2-10 shows also the relationship of the 100GBASE-R PCS sublayers (shown shaded) with other sublayers to the ISO OSI reference model.

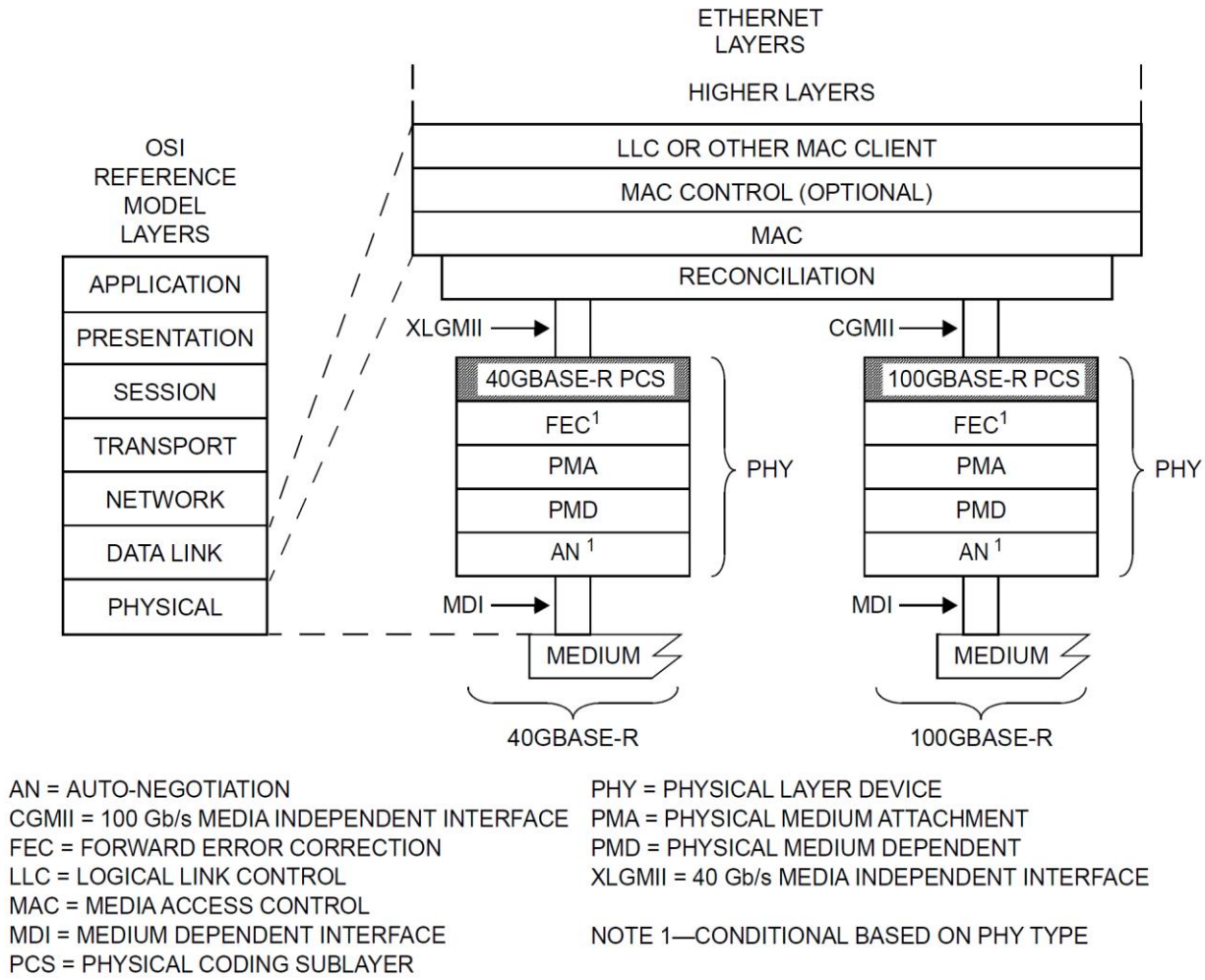


Figure 2-10. *100GBASE-R PCS relationship to the ISO/IEC OSI reference model and IEEE 802.3 Ethernet model.*

The PCS service interface is the MII (CGMII), which is defined in Section **2.3**. The 100Gb/s variant of this interface is called the 100Gb/s MII (CGMII). The CGMII provides a uniform interface to the RS for all 100Gb/s PHY implementations.

The 100GBASE-R PCSs provide all services required by the CGMII, including the following:

a. Encoding (decoding) of eight CGMII data octets to (from) 66-bit blocks (64B/66B).
b. Transferring encoded data to (from) the PMA.
c. Compensation for any rate differences caused by the insertion or deletion of alignment markers or due to any rate difference between the CGMII and PMA through the insertion or deletion of idle control characters.
d. Determining when a functional link has been established and informing the management entity via the MDIO when the PHY is ready for use.

The upper interface of the PCS may connect to the RS through the CGMII. The lower interface of the PCS connects to the PMA sublayer to support the PMD. If the optional BASE-R FEC sublayer is implemented and an optional physical instantiation, i.e., CAUI-n, is not implemented directly below the PCS sublayer, then the lower interface connects to the BASE-R FEC sublayer. For PHYs that use Section **2.8** RS-FEC, if an optional physical instantiation (i.e., CAUI-n) is not implemented directly below the PCS sublayer, then the lower interface connects to the RS-FEC sublayer. The 100GBASE-R PCS has a nominal rate at the PMA/FEC service interface of 5.15625Gtransfers/s per PCS lane, which provides capacity for the MAC data rate of 100Gb/s.

The PCS service interface allows the 100GBASE-R PCS to transfer information to and from a PCS client. The PCS client is the RS. The PCS Service Interface is precisely defined as the MII (CGMII) in Section **2.3**.

2.5 Physical Medium Attachment (PMA) sublayer, type 100GBASE-R

The PMA sublayer is common to the family of 100Gb/s PHY implementations, known as 100GBASE-R. The PMA allows the PCS (specified in Section **2.4**) to connect in a media-independent way with a range of physical media. The term 100GBASE-R is used when referring generally to PHYs using the PMA defined in this Section.

100GBASE-R can be extended to support any full duplex medium, which requires only that the PMD is compliant with the appropriate PMA interface.

The interfaces for the inputs of the 100GBASE-R PMAs are defined in an abstract manner and do not imply any particular implementation. The optional physical instantiation of the PMD service interfaces for 100GBASE-SR10 PMDs, known as CPPI, are defined in **Annex M**. For 100GBASE-R PMAs, electrical interfaces connecting PMA sublayers, known as CAUI-n, are defined in **Annex H**, **Annex I**, **Annex K**, and **Annex L**.

Figure 2-11 shows the relationship of the PMA sublayer (shown shaded) with other sublayers to the ISO OSI reference model.

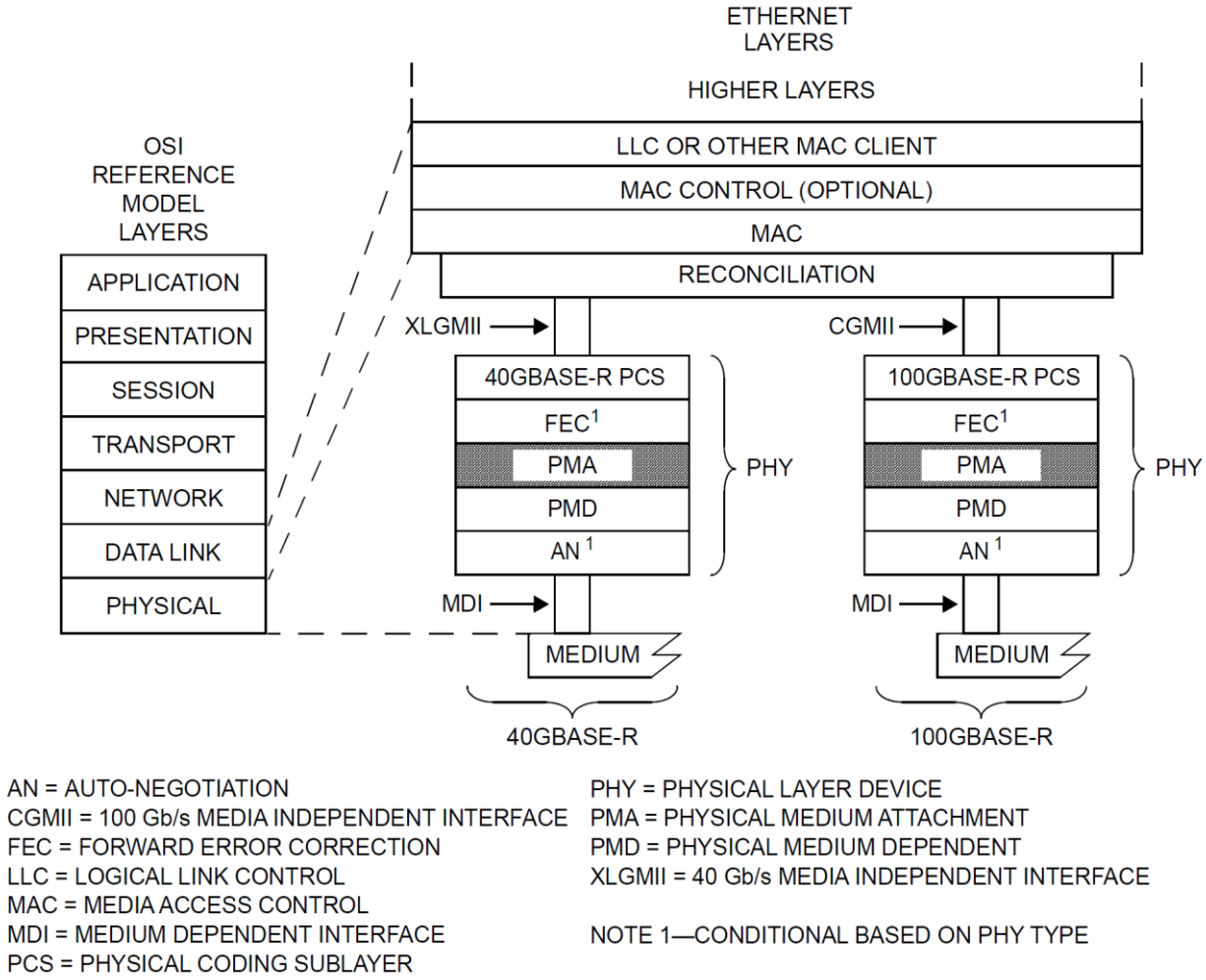


Figure 2-11. *40GBASE-R and 100GBASE-R PMA relationship to the ISO/IEC OSI reference model and IEEE 802.3 Ethernet model.*

The following summarizes the principal functions implemented (when required) by the PMA in both the transmit and receive directions:

a. Adapt the PCS Lane (PCSL) formatted signal to the appropriate number of abstract or physical lanes.
b. Provide per input-lane clock and data recovery.
c. Provide bit-level multiplexing.
d. Provide clock generation.
e. Provide signal drivers.
f. Optionally provide local loopback to/from the PMA service interface.
g. Optionally provide remote loopback to/from the PMD service interface.
h. Optionally provide test-pattern generation and detection.
i. Tolerate Skew Variation.

In addition, the PMA provides information about the receive link status in the receive direction.

An implementation can be use one or more PMA sublayers to adapt the number and rate of the PCS lanes to the number and rate of the PMD lanes. The number of the PMA sublayers required depends on the functionality partitioning for a particular implementation. An example is illustrated in **Figure 2-12**. This example illustrates the partitioning that might arise from use of an FEC device that is separate from the PCS. Additional examples are illustrated in **Annex J**. Each PMA maps the PCSLs from *p* PMA input lanes to *q* PMA output lanes in the Tx direction, and from *q* PMA input lanes to *p* PMA output lanes in the Rx direction.

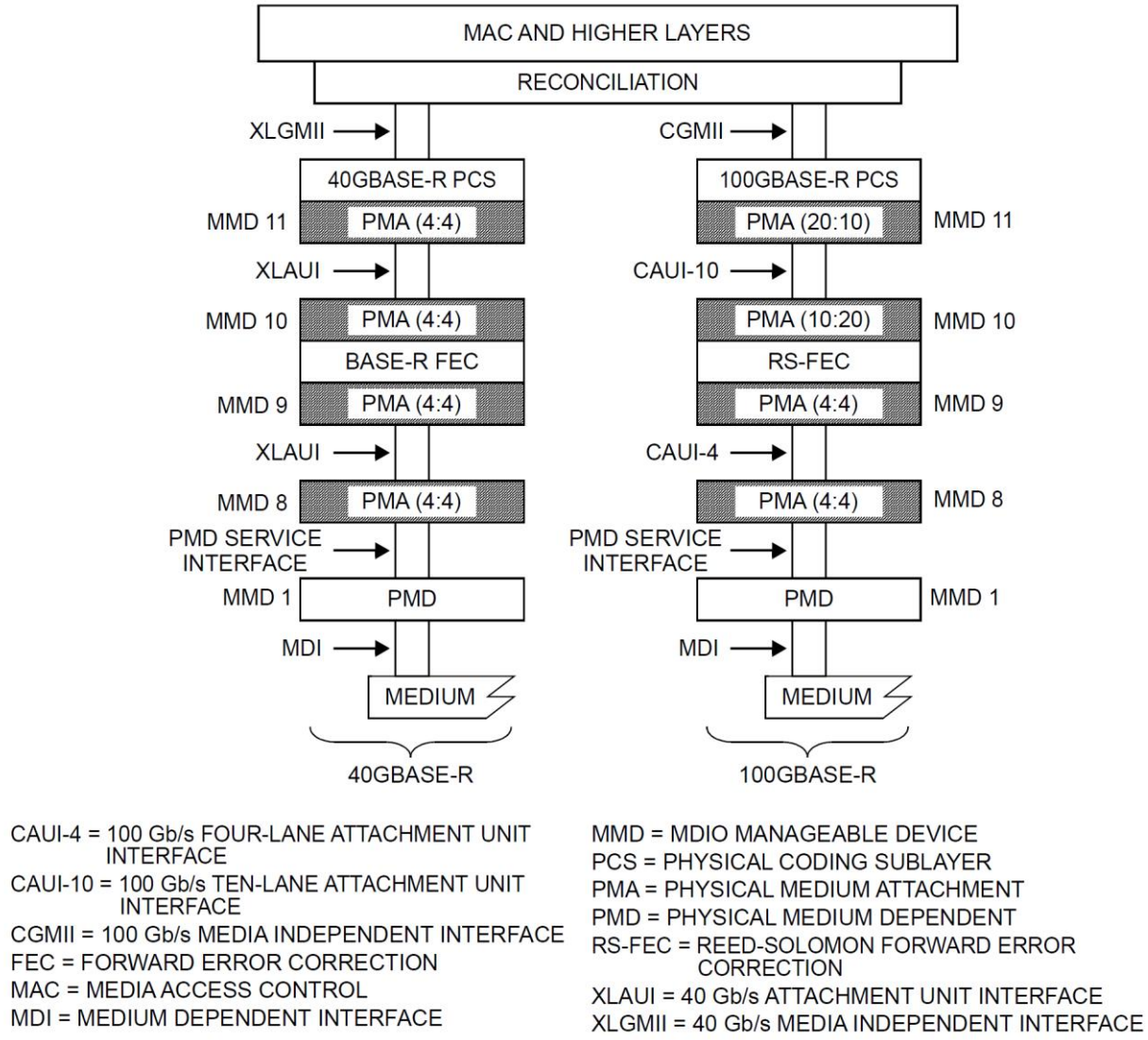


Figure 2-12. *Example 100GBASE-R PMA layering.*

Management Data Input/Output (MDIO) Manageable Device (MMD) addresses 1, 8, 9, 10, and 11 are available for addressing multiple instances of PMA sublayers (see

Table 2-6 for MMD device addresses). If the PMA sublayer which is closest to the PMD is packaged with the PMD, it shares MMD 1 with the PMD. If the PMD service interface is physically instantiated as nPPI (see **Annex M**), then the PMA sublayer which is closest to the PMD will be addressed as MMD 8. More addressable instances of PMA sublayers, each one separated from lower addressable instances by chip-to-chip interfaces, can be implemented and addressed allocating MMD addresses to PMAs in increasing numerical order going from the PMD toward the PCS. The example shown in **Figure 2-12** could be implemented with four addressable instances: MMD 8 addressing the lowest PMA sublayer (note that this cannot share MMD 1 with the PMD as they are not packaged together in this example), MMD 9 addressing the PMA sublayer above the CAUI-4 which is below the FEC, MMD 10 addressing the PMA sublayer below the CAUI-10 which is above the FEC, and MMD 11 addressing the PMA sublayer closest to the PCS.

Device address	MMD name
0	Reserved
1	PMA/PMD
2	WIS
3	PCS
4	PHY XS
5	DTE XS
6	TC
7	Auto-Negotiation
8	Separated PMA (1)
9	Separated PMA (2)
10	Separated PMA (3)
11	Separated PMA (4)
12	OFDM PMA/PMD
13	Power Unit
14 through 28	Reserved
29	RS and MII extension
30	Vendor specific 1
31	Vendor specific 2

Table 2-6. *MDIO Manageable Device addresses.*

The number of input and output lanes for a PMA is always divisors of the number of PCSLs. For PMA sublayers supporting 100GBASE-R PMDs, the number of PCSLs is 20.

The following guidelines apply to the partitioning of PMAs:

a. The inter-sublayer service interface is used for the PMA, FEC, and PMD service interfaces supporting a flexible architecture with optional FEC and multiple PMA sublayers.
    1. An instance of this interface can only connect service interfaces with the same number of lanes, where the lanes operate at the same rate.
b. CAUI-n are physical instantiations of the connection between two adjacent PMA sublayers.
    1. As a physical instantiation, it defines electrical and timing specification as well as requiring a receive re-timing function.
    2. CAUI-10 is a 10.3125GBd by 10 lane physical instantiation of the respective 100Gb/s connection.
    3. CAUI-4 is a 25.78125GBd by 4 lane physical instantiation of the respective 100Gb/s connection.
c. The abstract inter-sublayer service interface can be physically instantiated as a CAUI-n, using associated PMAs to map to the appropriate number of lanes.
d. Opportunities for optional test-pattern generation, optional test-pattern detection, optional local loopback and optional remote loopback are dependent upon the location of the PMA sublayer in the implementation.
e. A minimum of one PMA sublayer is required in a PHY.
f. A maximum of four PMA sublayers are addressable as MDIO MMDs.

2.6 Physical Medium Dependent (PMD) sublayer and medium, type 100GBASE-SR10

This Section specifies the 100GBASE-SR10 PMD together with the multimode fiber medium. When forming a complete PHY, a PMD shall be connected to the appropriate PMA as shown in **Table 2-7**, to the medium through the MDI, and optionally to the management functions that are accessible through the management interface, or equivalent.

Associated Section	100GBASE-SR10
2.3 - RS	Required
2.3 - CGMII [a]	Optional
2.4 - PCS for 100GBASE-R	Required
2.5 - PMA for 100GBASE-R10	Required
Annex H - CAUI-10 [b]	Optional
Annex I - Chip to module CAUI-10 [b]	Optional
Annex K - CAUI-4	Optional
Annex L - Chip-to-module CAUI-4	Optional
Annex M - CPPI	Optional
Annex F - Energy Efficient Ethernet	Optional
[a] CGMII are optional interfaces. If the appropriate interface is not implemented, a conforming implementation must functionally behave as the RS, and the CGMII were present. [b] If CAUI-n is present, there is at least a PMA between the CAUI-n and the PMD.	

Table 2-7. *Physical Layer Sections associated with the 100GBASE-SR10 PMDs.*

Figure 2-13 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC OSI reference model. 100Gb/s Ethernet is introduced in Section **2.2** and the purpose of each PHY sublayer is summarized in Section **2.3**.

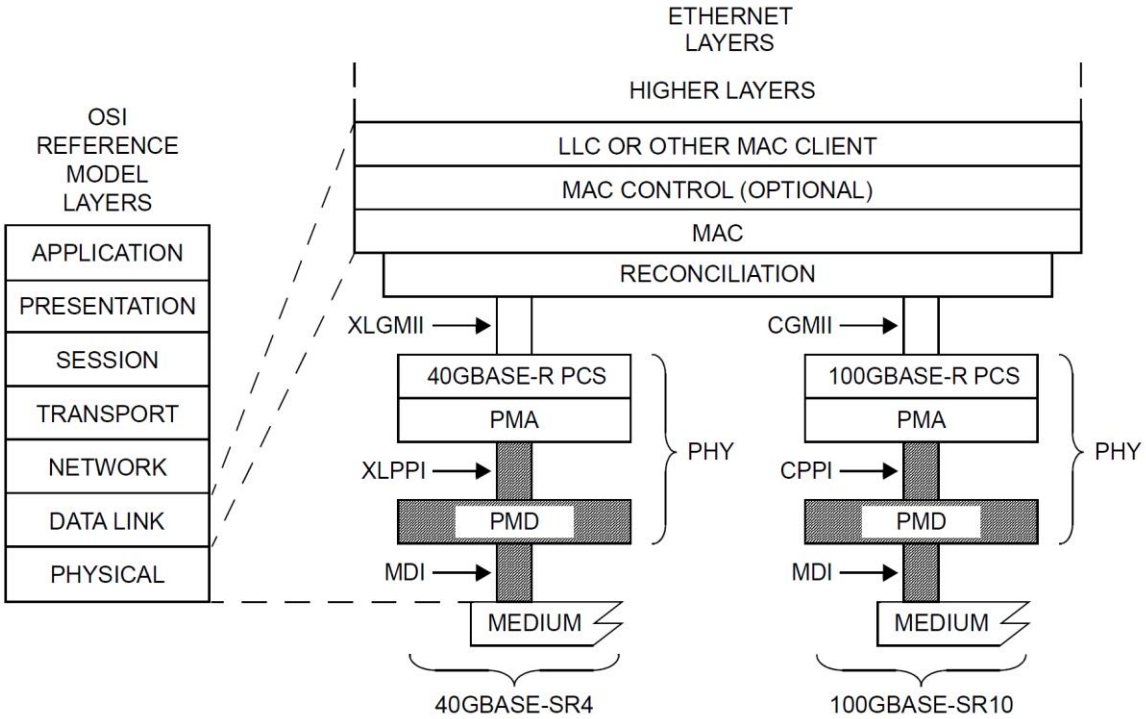


Figure 2-13. *40GBASE-SR4 and 100GBASE-SR10 PMDs relationship to the ISO/IEC OSI reference model and IEEE 802.3 Ethernet model.*

The 100GBASE-SR10 PMD sublayers provide point-to-point 100Gb/s Ethernet links over four or ten pairs of multimode fiber, up to at least 100m. **Table 2-8** shows the primary attributes of each PMD type.

PMD Type	100GBASE–SR10	Unit
Fiber type	50/125μm multimode, type A1a.2 [a] (OM3) or A1a.3 [b] (OM4)	
Number of fiber pairs	10	
Nominal wavelength	850	nm
Required operating range	0.5 to 100 for OM3	m
	0.5 to 150 for OM4 [c]	
Signaling rate, each lane	10.3125 ±100ppm	GBd
[a] Type A1a.2 (OM3) specified in IEC 60793-2-10. See Section **2.6.7**.		
[b] Type A1a.3 (OM4) specified in IEC 60793-2-10. See Section **2.6.7**.		
[c] This is an engineered link with maximum 1dB connection and splice loss.		

Table 2-8. *Summary of 100GBASE–SR10.*

100GBASE–SR10 uses ten of the same lanes. In this Section, where there are ten items (depending on PMD type) such as lanes, the items are numbered from 0 to n – 1, and an example item is numbered *i*. Thus, *n* is 10.

The connection to the PMA can use the optional physical instantiation of the PMD service interface called CPPI (ten lanes, for 100GBASE-SR10). The term “Parallel Physical Interface (nPPI)” is used to denote CPPI.

100GBASE-SR10 PHYs with the optional EEE capability can enter the fast wake LPI mode to conserve energy during periods of low link utilization (see **Annex F**). The EEE deep sleep mode is not supported.

The service interfaces for these PMDs are described in an abstract manner, although an optional implementation of the PMD service interface, the nPPI, is specified in **Annex M**. The PMD service interface supports the exchange of encoded data between the PMA entity that resides above the PMD, and the PMD entity. The PMD translates the encoded data to and from signals suitable for the specified medium.

The PMD service interface primitives are the following:

- PMD:IS_UNITDATA_i.request
- PMD:IS_UNITDATA_i.indication
- PMD:IS_SIGNAL.indication

The 100GBASE-SR10 PMD has ten parallel bit streams, where *i = 0* to *9* for 100GBASE-SR10.

The PMA (or the PMD) continuously sends four or ten parallel bit streams to the PMD (or the PMA), one per lane, each at a nominal signaling rate of 10.3125GBd.

Upon receipt of the *PMD:IS_UNITDATA_i.request* primitive, the PMD converts the specified streams of bits into the appropriate signals on the MDI.

The *PMD:IS_UNITDATA_i.indication* primitive corresponds to one of the signals received from the MDI. This primitive is received by the PMA.

The *PMD:IS_SIGNAL.indication* (SIGNAL_DETECT) primitive is generated by the PMD to report the parameter *SIGNAL_DETECT*, which indicates the status of the signals being received from the MDI (see Section **2.6.1.3**). There is one parameter and primitive, which reports on all the lanes as a group. This primitive is received by the PMA as *PMD:IS_SIGNAL.indication* (SIGNAL_OK).

NOTE: The *SIGNAL_DETECT = OK* does not guarantee that the *rx_bit* parameters are known to be good. It may refers to a poor-quality link, which is unable to provide sufficient light for a *SIGNAL_DETECT = OK* indication and still not meet the 10^{-12} BER objective.

The sum of the transmit and receive delays at one end of the link contributed by the 100GBASE-SR10 PMD including 2m of fiber in one direction shall be no more than 2048bit times (4 *pause_quanta* or 20.48ns). A description of the overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **2.2** and its references.

The Skew (relative delay) between the lanes must be kept within the appropriate limits, so that the information on the lanes can be reassembled by the PCS. The Skew Variation must also be limited to ensure that a given PCS lane always traverses the same physical lane. Skew and Skew Variation are defined in Section **2.2.2** and specified at the points SP1 to SP6 shown in **Figure 2-5** and **Figure 2-6**. Skew points as they relate to the nPPI are shown in **Figure 2-14**.

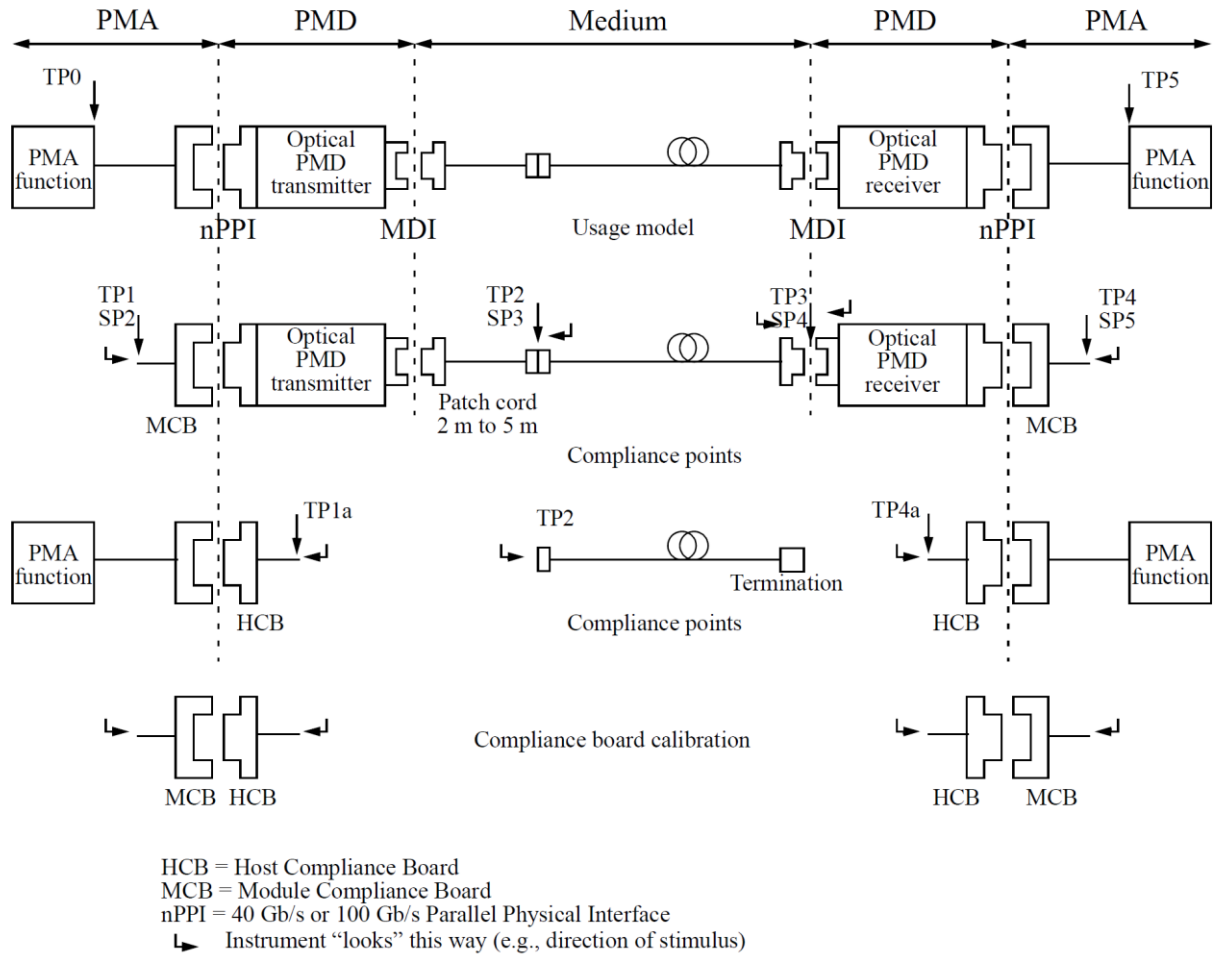


Figure 2-14. *Test points for 100GBASE-SR10.*

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

For more information on Skew and Skew Variation see Section **2.2.2**. The measurements of Skew and Skew Variation are defined in Section **2.6.4**.

2.6.1 PMD Functional Specifications

The 100GBASE–SR10 PMD perform the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

The PMD block diagram is shown in **Figure 2-15**. **Figure 2-14** shows the test points. It is not required that the PMD service interface be exposed or measurable (nPPI as defined in **Annex M** with compliance points TP1, TP1a, TP4, TP4a). However, a conforming implementation must behave as though the interface were compliant.

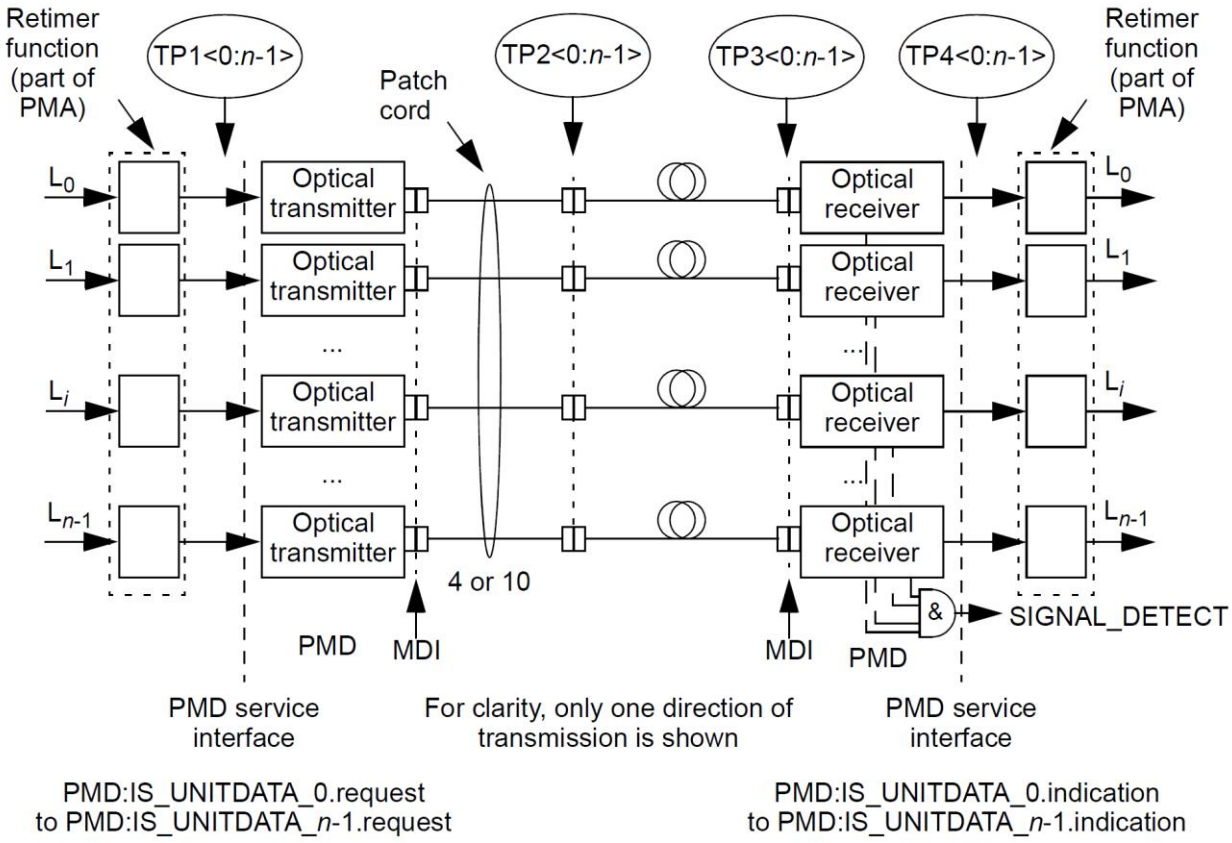


Figure 2-15. *Block diagram for 100GBASE–SR10 transmit/receive paths.*

For purposes of system conformance, the PMD sublayer is standardized at the test points described in Section **2.6.4**. The transmit side electrical signal (PMA output and PMD electrical input) is defined at TP1 and TP1a (see **Annex M**). The optical transmit signal is defined at the output end of a 50µm multimode fiber patch cord (TP2), between 2m and 5m in length. Unless otherwise specified, all optical transmitter tests and measurements defined in Section **2.6.4** are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see Section **2.6.7.1**). Unless specified otherwise, all optical receiver measurements and tests defined in Section **2.6.4** are made at TP3. The receive side electrical signal (PMD electrical output and PMA input) is defined at TP4 and TP4a (see **Annex M**).

2.6.1.1 PMD transmit function

The PMD Transmit function shall convert the four or ten electronic bit streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_n-1.request* into the same number of optical signal streams. The optical signal streams are delivered to the MDI, which contains ten parallel light paths

for transmit, according to the transmit optical specifications defined in this Chapter. The higher optical power level of each signal stream corresponds to *tx_bit = one*.

2.6.1.2 PMD receive function

The PMD Receive function converts the ten parallel optical signal streams received from the MDI into separate electronic bit streams for delivery to the PMD service interface using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_n-1.indication*, all according to the receive optical specifications in this Section. The higher optical power level of each signal stream corresponds to *rx_bit = one*.

2.6.1.3 PMD global signal detect function

The PMD global *SIGNAL_DETECT* function reports the state of *SIGNAL_DETECT* parameter via the PMD service interface. The *SIGNAL_DETECT* parameter is signaled continuously, while the *PMD:IS_SIGNAL.indication* message is generated when a change in *SIGNAL_DETECT* parameter occurs.

SIGNAL_DETECT is a global indicator of the presence of optical signals on all lanes. The value of the *SIGNAL_DETECT* parameter is generated according to the conditions defined in **Table 2-9**. The PMD receiver is not required to verify whether a compliant 100GBASE-SR10 signal is being received.

Receive conditions	SIGNAL_DETECT value
For any lane, Average optical power at TP3 ≤ -30dBm	FAIL
For all lanes, [(Optical power at TP3 ≥ Minimum OMA, each lane, in **Table 2-12**) and (Compliant 100GBASE–SR10 signal input as appropriate)]	OK
All other conditions	Unspecified

Table 2-9. *SIGNAL_DETECT value definition.*

Various implementations of the *SIGNAL_DETECT* function are permitted by this standard, including implementations that generate the values of the *SIGNAL_DETECT* parameter in response to the amplitude of the modulation of the optical signal and implementations, which respond to the average optical power of the modulated optical signal.

2.6.1.4 PMD lane-by-lane signal detect function

Various implementations of the *SIGNAL_DETECT* function are permitted by this standard, including implementations of MDIO, where each *PMD_signal_detect_i* (*i* represents the lane number in the range *0:n-1)* continuously sets in response to the optical signal on its associated lane, according to the requirements of **Table 2-9**.

2.6.2 Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for 100GBASE-SR10. While it is expected that a PMD will map electrical lane *i* to optical lane *i* and vice versa, there is no need to define the physical ordering of the lanes, as the PCS is capable of receiving the lanes in any arrangement. The positioning of transmit and receive lanes at the MDI is specified in Section **2.6.7.1**.

2.6.3 PMD to MDI specifications for 100GBASE-SR10

The required operating range for the 100GBASE–SR10 PMD is defined in **Table 2-8**. A compliant PMD operates on 50/125μm multimode fibers according to the specifications of **Table 2-10**. A PMD which exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., operating at 125m meets the operating range requirement of 0.5m to 100m). The signaling rate for a lane of a 100GBASE–SR10 PMD shall be as defined in **Table 2-8**. The optical signal at the transmit and receive side of the MDI is specified in Sections **2.6.3.1** and **2.6.3.3**. The range of optical signals within the optical medium is defined in Section **2.6.3.2**, and an illustrative link power budget is provided in Section **2.6.3.4**. Test points are defined in Section **2.6.4**.

Description	OM3 [a]	OM4 [b]	Unit
Nominal core diameter	50		μm
Nominal fiber specification wavelength	850		nm
Effective modal bandwidth (min) [c]	2000	4700	MHz•km
Cabled optical fiber attenuation (max)	3.5		dB/km
Zero dispersion wavelength (λ_0)	$1295 \le \lambda_0 \le 1340$		nm
Chromatic dispersion slope (max) (S_0)	0.105 for $1295 \le \lambda_0 \le 1310$ and $0.000375 \times (1590 - \lambda_0)$ for $1310 \le \lambda_0 \le 1340$		ps/nm^2 km
[a] IEC 60793-2-10 type A1a.2.			

[b] IEC 60793-2-10 type A1a.3.

[c] When measured with the launch conditions specified in **Table 2-11**.

***Table 2-10**. Optical fiber and cable characteristics.*

2.6.3.1 Transmitter optical specifications

Each lane of a 100GBASE–SR10 optical transmitter shall meet the specifications of **Table 2-11** per the definitions in Section **2.6.4**.

Description	Type	Value	Unit
Center wavelength	Range	840 to 860	nm
RMS spectral width [a]	Max	0.65	nm
Average launch power, each lane	Max	2.4	dBm
Average launch power, each lane	Min	-7.6	dBm
Optical Modulation Amplitude (OMA), each lane	Max	3	dBm
OMA, each lane	Min	-5.6 [b]	dBm
Difference in launch power between any two lanes (OMA)	Max	4	dB
Peak power, each lane	Max	4	dBm
Launch power in OMA minus TDP, each lane	Min	-6.5	dBm
Transmitter and dispersion penalty (TDP), each lane	Max	3.5	dB
Extinction ratio	Min	3	dB
Optical return loss tolerance	Max	12	dB
Encircled flux [c]		≥ 86% at 19μm, ≥ 30% at 4.5μm	
Transmitter eye mask definition (X1, X2, X3, Y1, Y2, Y3) Hit ratio 5×10^{-5} hits per sample	Spec values	0.23, 0.34, 0.43, 0.27, 0.35, 0.4	
Average launch power of OFF transmitter, each lane	Max	-30	dBm

[a] RMS spectral width is the standard deviation of the spectrum.

[b] Even if the TDP < 0.9dB, the OMA (min) must exceed this value.

[c] If measured 50μm fiber type A1a.2 or type A1a.3, in accordance with IEC 61280-1-4.

***Table 2-11.** 100GBASE–SR10 optical transmit characteristics.*

2.6.3.2 Characteristics of signal within, and at the receiving end of, a compliant optical channel

Table 2-12 gives the characteristics of a signal within, and at the receiving end of, a lane of a compliant 100GBASE–SR10 optical channel, and the aggregate signal. A

signal with power in OMA or average power not within the ranges given is not compliant.

Description	Minimum		Maximum	Unit
Fiber type	OM3	OM4		
Total average power for 100GBASE–SR10	+0.5	+0.9	+12.4	dBm
Average power, each lane	-9.5	-9.1	+2.4	dBm
OMA, each lane	-7.5	-7.1	+3	dBm

Table 2-12. *Characteristics of signal within, and at the receiving end of, a compliant optical channel*

2.6.3.3 100GBASE-SR10 receiver optical specifications

Each lane of a 100GBASE–SR10 optical receiver shall meet the specifications defined in **Table 2-13** per the definitions in Section **2.6.4**.

Description	Type	Value	Unit
Center wavelength, each lane	Range	840 to 860	nm
Damage threshold [a]	Min	+3.4	dBm
Average power at receiver input, each lane	Max	+2.4	dBm
	Min	-9.5	dBm
Receiver reflectance	Max	-12	dB
OMA, each lane	Max	3	dBm
Stressed receiver sensitivity in OMA, each lane [b]	Max	-5.4	dBm
Peak power, each lane	Max	4	dBm
Conditions of stressed receiver sensitivity test:			
- Vertical eye closure penalty (VECP) [c], each lane	-	1.9	dB
- Stressed eye J2 Jitter [c], each lane	-	0.3	UI
- Stressed eye J9 Jitter [c], each lane	-	0.47	UI
- OMA of each aggressor lane	-	-0.4	dBm
Receiver jitter tolerance in OMA, each lane [d]	Max	-5.4	dBm
Conditions of receiver jitter tolerance test:			
- Jitter frequency and peak-to-peak amplitude	-	(75, 5)	(kHz, UI)
- Jitter frequency and peak-to-peak amplitude	-	(375, 1)	(kHz, UI)
- OMA of each aggressor lane	-	-0.4	dBm

[a] The receiver shall be able to tolerate without damage, and continuously exposed to a modulated optical input signal of this power level on one lane.

[b] Measured with conformance test signal at TP3 (see Section **2.6.4.9**).

[c] Vertical eye closure penalty and stressed eye jitter are test conditions for measuring stressed receiver sensitivity. They are not characteristics of the receiver. The apparent discrepancy between the VECP and the TDP is because the VECP is defined at the eye center, while the TDP is defined with ±0.15 UI offsets of the sampling instant.

[d] This is a test of the optical receiver's ability to track low-frequency jitter and is inappropriate for any subsystem that does not include a CRU.

Table 2-13. *100GBASE-SR10 optical receiver characteristics.*

2.6.3.4 100GBASE-SR10 illustrative link power budget

Illustrative power budgets and penalties for 100GBASE–SR10 optical channels are shown in **Table 2-14**.

Parameter	OM3	OM4	Unit
Effective modal bandwidth at 850 nm [a]	2000	4700	MHz•km
Power budget (for maximum TDP)	8.3		dB
Operating distance	0.5 to 100	0.5 to 150	m
Channel insertion loss [b]	1.9	1.5	dB
Allocation for penalties (for maximum TDP) [c]	6.4	6.5	dB
Unallocated margin	0	0.3 [d]	dB
Additional insertion loss allowed	0		dB

[a] Per IEC 60793-2-10.

[b] The channel insertion loss is calculated using the maximum distances specified in **Table 2-8** and cabled optical fiber attenuation of 3.5dB/km at 850nm plus an allocation for connection and splice loss given in Section **2.6.7**.

[c] Link penalties are used for link budget calculations.

[d] This unallocated margin is not available for use.

Table 2-14. *40GBASE–SR4 or 100GBASE–SR10 illustrative link power budgets.*

2.6.4 Definitions of optical and dual-use parameters and measurement methods

In this section, the Test Points (TPs) and the parameters applicable to optical signals and for dual use (both optical and electrical) are defined.

Figure 2-14 shows the six test points for 100GBASE-SR10. These TPs are TP1, TP1a, TP2, TP3, TP4, and TP4a, where four of these are Skew Points (SPs) SP2, SP3, SP4, and SP5 as shown. **Figure 2-14** also shows the substitution of compliance boards for PMD or PMA. This is explained in **Annex M**. **Table 2-15** shows the parameters or signals measured at each point, including the electrical compliance points.

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified. A patch cord that connects the MDI transmit side to 4 or 10 individual connectors may be suitable.

Test point	Direction	Parameter
TP1	Looking downstream into PMD transmitter input	PMD transmitter input return loss
TP1a	Looking upstream into PMA transmitter output	PMA transmitter output signal and output return loss, PMD transmitter compliance signal calibration, PMA receiver compliance crosstalk signal calibration
TP2	Looking upstream into optical transmitter output	PMD transmitted signal, PMD transmitter reflectance
	Looking downstream into fiber	Optical return loss, connector reflections
TP3	Looking upstream into fiber	Looking upstream into fiber
	Looking downstream into optical receiver input	PMD receiver reflectance
TP4	Looking upstream into PMD receiver output	PMD receiver output signal and output return loss, PMA receiver compliance signal calibration
TP4a	Looking downstream into PMA receiver input	PMA receiver input return loss

Table 2-15. *Parameters defined at each test point.*

Compliance is achieved in normal operation, so that specific test patterns must be defined for measurement consistency and for the measurement of some parameters. **Table 2-16** lists the defined test patterns, and **Table 2-17** gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subsections in which each parameter is defined. As Pattern 3 is more demanding

than Pattern 5 (which itself is the same or more demanding than other 100GBASE-R bit streams) an item that is compliant using Pattern 5 is considered as compliant even if it does not comply with the required limit using Pattern 3. Test patterns for further electrical parameters are given in **Table M. 6**.

Pattern no.	Pattern
Square wave	Square wave (8 ones, 8 zeros)
3	PRBS31
4	PRBS9
5	Scrambled idle

Table 2-16. *Test patterns.*

Parameter	Pattern	Related Section
Wavelength, spectral width	3, 5, or valid 100GBASE-SR10 signal	**2.6.4.1**
Average optical power	3, 5, or valid 100GBASE-SR10 signal	**2.6.4.2**
Transmitter OMA (modulated optical power)	Square wave or 4	**2.6.4.3**
Extinction ratio	3, 5, or valid 100GBASE-SR10 signal	**2.6.4.6**
Transmitted optical waveform (eye mask)	3, 5, or valid 100GBASE-SR10 signal	**2.6.4.9, 2.6.4.7**
TDP (transmitter and dispersion penalty)	3 or 5	**2.6.4.4**
Stressed receiver sensitivity	3 or 5	**2.6.4.9**
Receiver jitter tolerance	3 or 5	**2.6.4.10**
Calibration of OMA for receiver tests	Square wave or 4	
Vertical eye closure penalty calibration	3 or 5	
J2 Jitter	3, 5, or valid 100GBASE-SR10 signal	**2.6.4.9**
J9 Jitter	3 or 5	**2.6.4.9**
Data Dependent Pulse Width Shrinkage (DDPWS)	4	**Annex M**
AC common-mode voltage	3, 5, or valid 100GBASE-SR10 signal	**Annex M**
Transition time	Square wave or 4	**Annex M**
Electrical waveform (eye mask)	3, 5, or valid 100GBASE-SR10 signal	**2.6.4.9, Annex M**

Table 2-17. *Test patterns and related subsections.*

TDP is defined for each lane, at a BER of 10^{-12} on that lane. Stressed receiver sensitivity, receiver jitter tolerance and host input signal tolerance (in **Annex M**) are

defined for an interface BER of 10^{-12}. The interface BER is the average of the ten receive lanes BER when they are stressed.

Measurements with Pattern 3 (PRBS31) are defined for the lane-by-lane BER measurements. Measurements with Pattern 5 (scrambled idle) provide the interface BER, if all lanes are stressed at the same time. In turn, if each lane is stressed, the BER is diluted by the nine unstressed lanes, and the BER for that stressed lane alone must be found, e.g., by multiplying by 10 if the unstressed lanes have low BER. To provide the TDP measurement with Pattern 5, unstressed lanes for the error detector must be created by setting the power at the reference receivers above their sensitivities, or by copying the contents of the transmit lanes not under the BER test to the error detector by other means. In stressed receiver sensitivity and receiver jitter tolerance measurements, unstressed lanes may be created by setting the power at the receiver under test well above its sensitivity and/or not stressing those lanes with InterSymbol Interference (ISI) and jitter, or by other means. Each receive lane is stressed in turn, while all are operated. All aggressor lanes are operated as specified. The interface BER is defined as the average BER of all the lanes when stressed.

Where relevant, the parameters are defined with all co-propagating and counter-propagating lanes operational, so that crosstalk effects are included. Where not otherwise specified, the maximum amplitude (OMA or Voltage Modulation Amplitude (VMA)) for a particular situation is used, and for the counter-propagating lanes, the minimum transition time is used. In any other case, alternative test methods that generate equivalent results must be used. While the lanes in a particular direction can share a common clock, the Tx and Rx directions are not synchronous to each other. If Pattern 3 is used for the lanes not under test using a common clock, there is at least 31 UI delay between the PRBS31 patterns on one and any other lane.

2.6.4.1 Wavelength and spectral width

The wavelength and spectral width of each optical lane shall be within the range given in **Table 2-11** if measured using the method given in the centroidal wavelength and RMS spectral width definitions in IEC 61280-1-3. The lane under test is modulated using the test pattern defined in **Table 2-17**.

2.6.4.2 Average optical power

Average optical power shall comply with the methods given in IEC 61280-1-1.

2.6.4.3 Optical Modulation Amplitude

In this Section is defined OMA for measurement with a square wave (8 ones, 8 zeros) test pattern and the variable *MeasuredOMA* for measurement with a PRBS9 test pattern, where each optical lane shall be tested individually.

OMA is the difference in optical power for the nominal “1” and “0” levels of the optical signal, OMA is shown in **Figure 2-16**, using waveform averaging or histogram means.

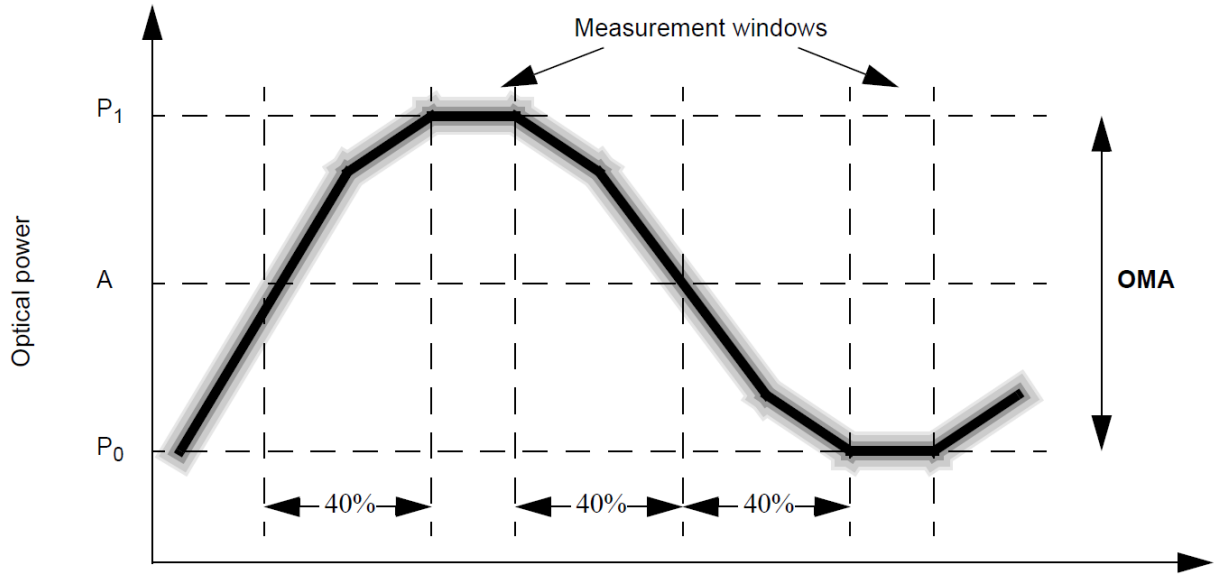


Figure 2-16. *Optical modulation amplitude waveform measurement.*

Compliance is to be achieved in normal operation. Two types of test pattern are used, square wave (a pattern consisting of four to eleven consecutive ones followed by an equal run of zeros can be used as a square wave. These patterns have fundamental frequencies between approximately 452MHz and 1289MHz) and other based on 10GBase-R or 10GBase-W (Patterns 1, 2, and 3 are defined in **Table 2-18**. Pattern 3 is optional.).

Pattern	Pattern for 10GBASE-R	Pattern for 10GBASE-W
1	$B_n\ B_i\ B_n\ B_i$	Mixed
2	$A_n\ A_i\ A_n\ A_i$	
3	PRBS31	PRBS31

Table 2-18. *Test patterns.*

For 10GBASE-R, two test pattern segments are specified, in two variants, the “normal” (*n*) and the “inverted” (*i*). Both are 8448 bits long. They can be generated dynamically by the 58 bit scrambler and “control block” sync header generation and

using the scrambler starting seeds specified in **Table 2-19**. The segments are assembled into patterns, each containing four segments, as described in **Table 2-18**. Each pattern can be held as a static one in test equipment or generated or detected dynamically.

Segments	Seed[57:0] [a]
A_n	0x3C8B44DCAB6804F
B_n	0x34906BB85A38884
[a] The "invert" segments A_i and B_i are generated using the inverted seeds for A_n and B_n, respectively.	

Table 2-19. *Pattern segments.*

Each segment contains a sync header transition every 66 bits.

The recommended technique for measuring optical modulation amplitude is shown in **Figure 2-17**. Optionally, a fourth-order Bessel-Thomson filter can be used after the O/E converter.

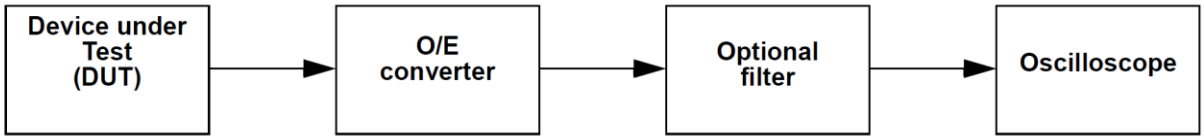


Figure 2-17. *Recommended technique for measurement of optical amplitude modulation.*

The test regime consisting of the O/E converter, the optional filter and the oscilloscope has the following requirements:

a. The bandwidth shall be at least 3/T, where *T* is the time at high or low (00001111 giving approximately *T* = 400ps and 7.5GHz as an example); and
b. The measurement shall be calibrated at the appropriate wavelength for the transmitter under test.

With the Device Under Test (DUT) transmitting the square wave described above, the following procedure shall be used to measure optical amplitude modulation:

a. Configure the test equipment as shown in **Figure 2-17**.
b. Measure the mean optical power P_1 of the logic "1" as defined over the center 20% of the time interval where the signal is in the high state. (See **Figure 2-16**.)
c. Measure the mean optical power P_0 of the logic "0" as defined over the center 20% of the time interval where the signal is in the low state. (See **Figure 2-16**.)
d. $OMA = P_1 - P_0$.

A method of approximating OMA is shown in **Figure 2-18**.

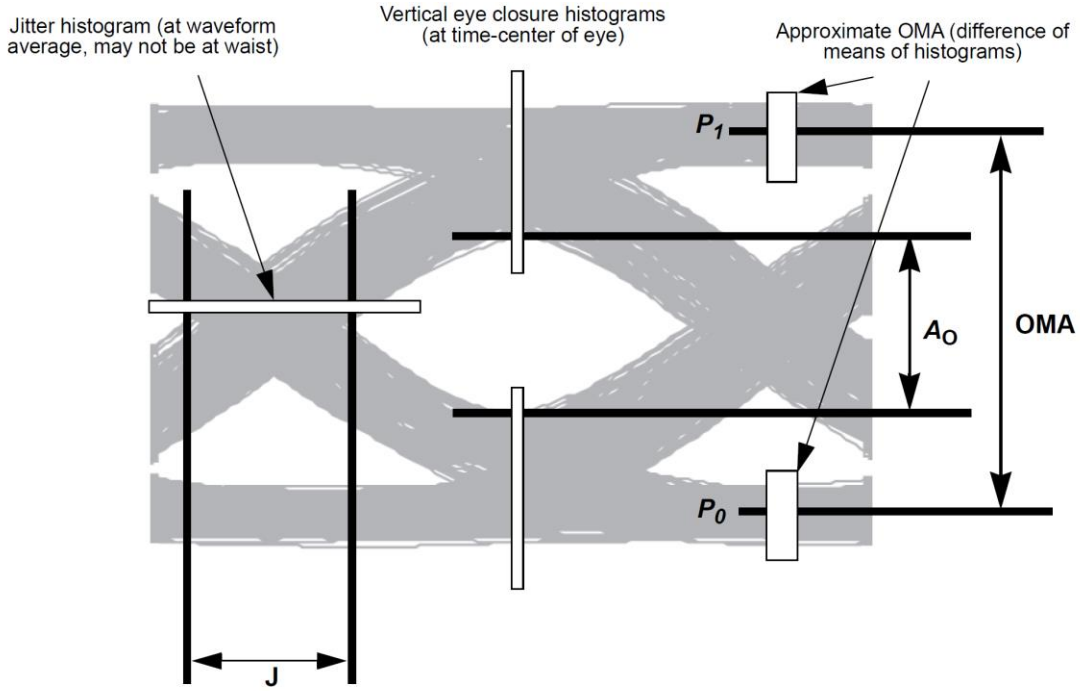


Figure 2-18. *Required characteristics of the test signal at TP3.*

OMA is also defined by the measurement method as illustrated in **Figure 2-19**. The mean logic ONE and mean logic ZERO values are measured over the center 20% of the two-time intervals of the square wave. The OMA is the difference between these two means.

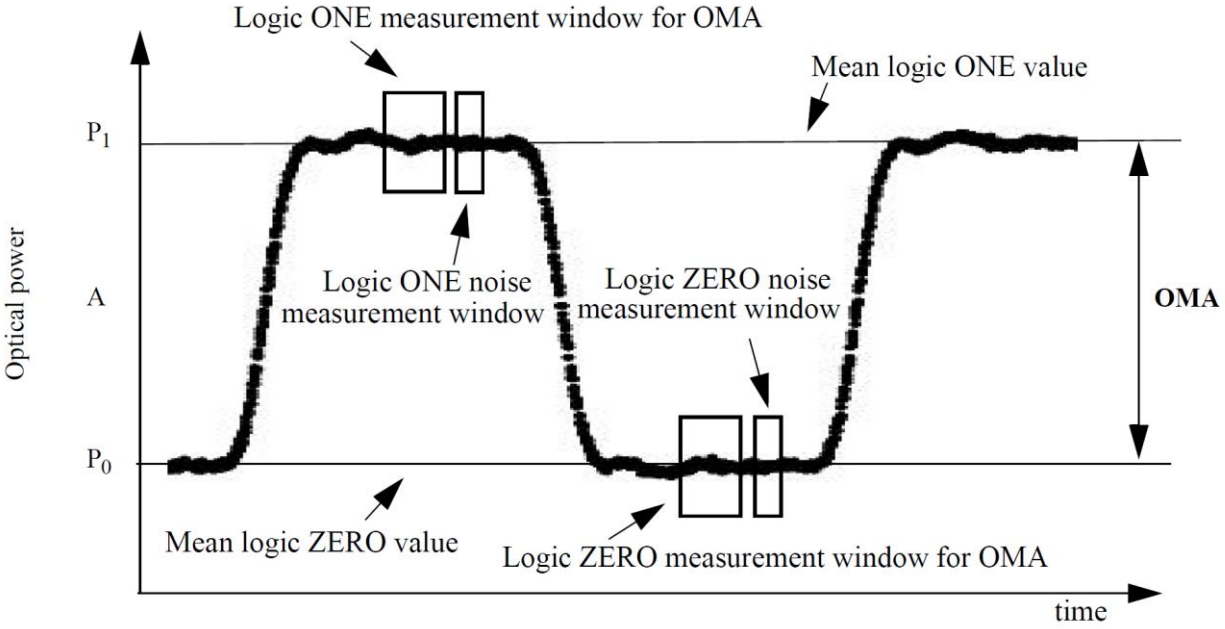


Figure 2-19. *Positions of logic ZERO and logic ONE measurement windows for OMA and transmitter noise measurements.*

2.6.4.4 Transmitter and dispersion penalty

The Transmitter and Dispersion Penalty (TDP) measurements are specified for the transmitter impairments with chromatic effects for a transmitter to be used with single-mode fiber, and for the transmitter impairments with modal (not chromatic) dispersion effects for a transmitter to be used with multimode fiber.

The setup for measurement of transmitter and dispersion penalty is shown in **Figure 2-20** and consists of a reference transmitter, the transmitter under test, a controlled optical reflection, an optical attenuator, a test fiber, a reference receiver, a transversal filter for 10GBASE-S, and a bit-error ratio tester. For 10GBASE-S, the polarization rotator shown in **Figure 2-20** is removed from the setup and the single-mode fiber replaced with a multimode fiber. All BER and sensitivity measurements are made with the test patterns in Section **2.6.4.3**.

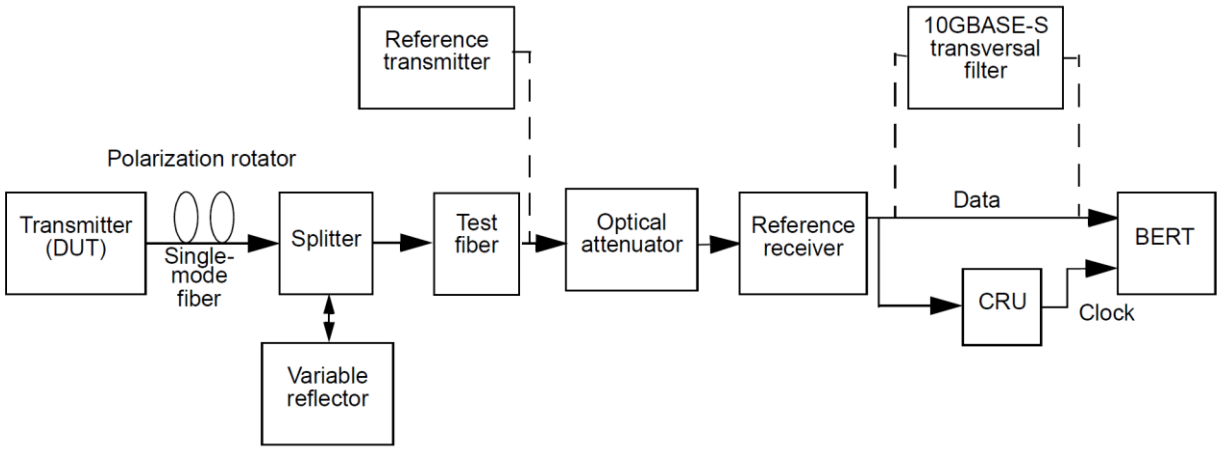


Figure 2-20. *Test setup for the measurement of a transmitter test and for the dispersion penalty.*

TDP is defined for 10GBASE-S with the following exceptions:

a. Each optical lane is tested individually with all other lanes in operation.
b. The test pattern is as defined in **Table 2-17**.
c. The transmitter is tested using an optical channel with an optical return loss of 12dB.
d. The reference receiver has a fourth order Bessel-Thomson filter response with a bandwidth of 6.1GHz. The transversal filter is not used.
e. The reference sensitivity S and the measurement *P_DUT* are both measured with the sampling instant displaced from the eye center by ± 0.15 UI.
f. The test setup illustrated in **Figure 2-20** shows the reference method.

g. The BER of 10^{-12} is for the lane under test on its own. See Section **2.6.4** for multi-lane pattern considerations.

2.6.4.5 Relative intensity noise optical modulation amplitude measuring procedure

The Relative Intensity Noise Optical Modulation Amplitude (RIN_xOMA) measuring procedure describes a component test, which cannot be appropriate for a system level test depending on the implementation.

The test arrangement is shown in **Figure 2-21**. The optical path between the DUT and the detector has a single discrete reflection with the specified optical return loss as seen by the DUT.

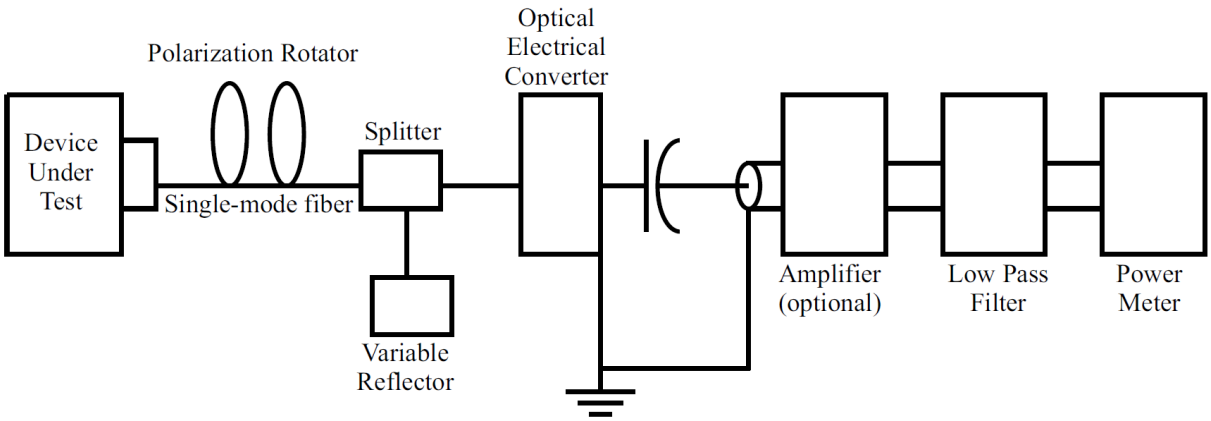


Figure 2-21. *RIN_xOMA measurement setup.*

Both the OMA power and the noise power are measured by AC-coupling the O/E converter into the high frequency electrical power meter. If necessary, an amplifier can be used to boost the signal to the power meter. A low-pass filter is used between the photodetector and the power meter to limit the noise measured to the passband, which is appropriate to the data rate of interest. In order to measure the noise, the modulation to the Device Under Test (DUT) is turned off.

The optical path and detector combination can be configured for a single dominant reflection with an optical return loss. The length of the test cable is not critical but should be in excess of 2m.

The polarization rotator is required for the transformation of an arbitrary orientation elliptically polarized wave into a fixed orientation linearly polarized wave.

If necessary, the noise can be amplified to a level consistent with accurate measurement by the power meter.

The upper -3dB limit of the measurement apparatus must be approximately equal to the bit rate (i.e., 10GHz). The total filter bandwidth used in the RIN calculation is required in order to take into consideration the low-frequency cutoff of the DC-blocking capacitor. The low-frequency cutoff is recommended to be less than 1MHz.

The low pass filter must be placed in the circuit as the last component before the power meter to eliminate any high-frequency noise components generated by the detector/amplifier.

The RMS electrical power meter must be zero in the absence of input optical power to remove any residual noise.

Use the following procedure to test relative intensity noise optical modulation amplitude:

a. with the DUT disconnected, zero the power meter, and connect the DUT,
b. turn on the laser, and ensure that the laser is not modulated,
c. operate the polarization rotator, while observing the power meter output to maximize the noise read by the power meter (note the maximum power, P_N),
d. turn on the modulation to the laser using the square wave pattern and note the power measurement, P_M,
e. calculate RIN from the observed electrical signal power and noise power by use of **Equation 2-2**.

$$RIN_xOMA = 10log\frac{P_N}{BW \times P_M}(dB/Hz)$$

Equation 2-2

where:

RIN_xOMA is the Relative Intensity Noise referred to optical modulation amplitude measured with x dB reflection,

P_N is the electrical noise power in watts with modulation off,

P_M is electrical power in watts with modulation on,

BW is the low-pass bandwidth of filter – high-pass bandwidth of DC-blocking capacitor [noise bandwidth] of the measuring system (Hz).

For testing multimode components or systems, the polarization rotator must be removed from the setup and the single-mode fiber must be replaced with a multimode fiber.

2.6.4.6 Extinction ratio

Extinction ratio shall comply with the methods given in IEC 61280–2–2 using the test pattern defined in **Table 2-17**.

NOTE: Extinction ratio and OMA are defined with different test patterns (see **Table 2-17**).

2.6.4.7 Transmitter optical waveform (transmit eye)

The required optical transmitter pulse shape characteristics are specified in the form of a mask of the transmitter eye diagram specified in **Table 2-11** and shown in **Figure 2-22**. The transmitter optical waveform of a port transmitting the test pattern specified in **Table 2-17** shall meet specifications according to the methods specified in Section **2.6.4.8** with the filter nominal reference frequency f_r of 7.5GHz and the filter tolerances as specified for STM-64 in ITU-T G.691.

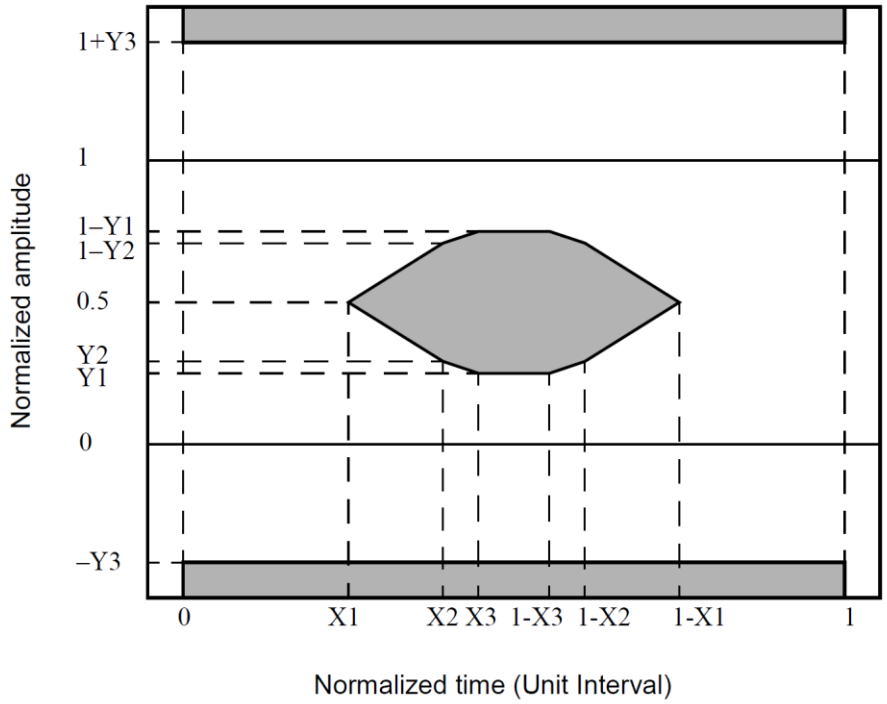


Figure 2-22. *Transmitter eye mask definition.*

2.6.4.8 Optical transmitter eye mask

The eye is measured with respect to the mask specified in **Table 2-11** and shown in **Figure 2-22** using a receiver with the fourth-order Bessel-Thomson response having a transfer function given by **Equation 2-3** and **Equation 2-4**:

$$H(y) = \frac{105}{105 + 105y + 45y^2 + 10y^3 + y^4}$$

Equation 2-3

where

$$y = 2.114p, \quad p=\frac{j\omega}{\omega_r}, \quad \omega_{r=2\pi f_r}, \quad f_r = Reference\ frequency\ in\ GHz$$

Equation 2-4

The Bessel-Thomson receiver is intended to provide uniform measurement conditions at the transmitter. Compensation can be made for the variation of the reference receiver filter response from the ideal fourth-order Bessel-Thomson response.

The normalized levels of 0 and 1 represent the logic zero and the one, respectively. These levels are defined by the means of the lower and upper halves of the central 0.2 UI of the eye.

The transmitter shall achieve a hit ratio lower than the limit of hits per sample specified in the appropriate table or 5×10^{-5} hits per sample if not otherwise specified. “Hits” are the number of samples within the grey areas of **Figure 2-22**, and the sample count is the total number of samples from 0 UI to 1 UI.

The optical eye pattern measurement procedures shall comply with the ones found in IEC 61280-2-2.

2.6.4.9 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in **Table 2-13** with the conformance test signal at TP3 and with the following exceptions:

a. The reference test procedure for a single lane.
b. The sinusoidal amplitude interferer is replaced by a Gaussian noise generator.
c. The fourth-order Bessel-Thomson filter is replaced by a low-pass filter followed by a limiter and a fourth-order Bessel-Thomson filter.
d. The sinusoidal jitter is at a fixed 80MHz frequency and between 0 and 0.05 UI peak-to-peak amplitude.
e. The Gaussian noise generator, the amplitude of the sinusoidal jitter, and the Bessel-Thomson filter are adjusted so that the VECP, J2 Jitter and J9 Jitter specifications given in **Table 2-13** are simultaneously met (the random noise effects such as RIN, random clock jitter do not need to be minimized).
f. The pattern for the received compliance signal is specified in **Table 2-17**.
g. The interface BER of the PMD receiver is the average of the BER of all receive lanes while stressed and at the specified receive OMA.

h. Where nPPI CAUI-n is exposed, a PMD receiver is considered compliant if it meets the module electrical output specifications at TP4 given in **Table M. 3** for nPPI, or the requirements in **Table I. 3** for CAUI-10, or the requirements in **Table L. 3** for CAUI-4.

Stressed receiver sensitivity is defined with the transmit and receive lanes in operation. Each receive lane is tested in turn while all aggressor receive lanes are operated as specified in **Table 2-13**. Pattern 3 or Pattern 5, or a valid 100GBASE–SR10 signal is sent from the transmit section of the receiver under test. The signal being transmitted is asynchronous to the received signal.

For 100GBASE-SR10, the relevant BER is the interface BER. The interface BER is the average of the four or ten BERs of the receive lanes when stressed.

2.6.4.10 Receiver jitter tolerance

Receiver jitter tolerance shall be with the following differences:

a. The reference test procedure for a single lane.
b. The pattern to be received is specified in **Table 2-17**.
c. The parameters of the signal are specified in **Table 2-13** and the power in OMA at the receiver is set to the maximum for receiver jitter tolerance in OMA given in **Table 2-13**.
d. Each receive lane is tested in turn while all are operated. All aggressor lanes are operated as specified in **Table 2-13**.
e. The receive lanes not being tested are receiving Pattern 3, Pattern 5, or a valid 100GBASE-SR10 signal.
f. The transmitter is transmitting one of these signals using all lanes.
g. The transmitter and the receiver are not synchronous.
h. The interface BER of the PMD receiver is the average of the BER of all receive lanes when stressed.
i. The mode-conditioning patch cord suitable for 62.5/125μm fiber is not used.

2.6.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

100GBASE–SR10 optical transceivers shall conform to Hazard Level 1M laser requirements as defined in IEC 60825–1 and IEC 60825–2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Laser safety standards and regulations require that the manufacturer of a laser product shall provide information about the product's laser, safety features, labeling, use, maintenance, and service.

It is recommended that proper installation practices, as defined by applicable local codes and regulation, shall be followed.

The fiber optic cabling (channel) contains 4 or 10 optical fibers for each direction to support 100GBASE-SR10. The fiber optic cabling connects the transmitters at the MDI on one end of the channel to the receivers at the MDI on the other end of the channel. As defined in Section **2.6.7.1**, the optical lanes appear in defined locations at the MDI, but the locations are not assigned specific lane numbers, because any transmitter lane can be connected to any receiver lane. The fiber optic cabling model (channel) is the same as a simplex fiber optic link segment. The term channel is used in consistency with the generic cabling standards.

2.6.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 2-23**.

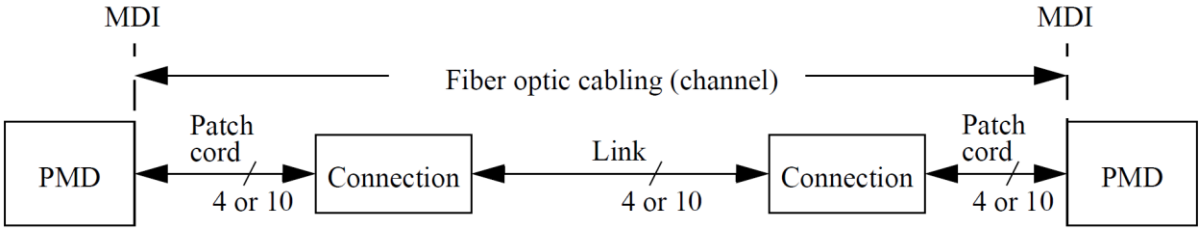


Figure 2-23. *Fiber optic cabling model.*

The channel shall comply with the specifications in **Table 2-20**.

Description	Type	OM3	OM4	Unit
Operating distance	Max	100	150	m
Cabling Skew	Max	79		ns
Cabling Skew Variation [a]	Max	2.2		ns
Channel insertion loss	Min	0		dB
Channel insertion loss [b]	Max	1.9 [c]	1.5 [d]	dB

[a] An additional 0.6ns of Skew Variation could be caused by wavelength changes, which are attributable to the transmitter not the channel.

[b] These channel insertion loss values include cable, connectors, and splices.

[c] 1.5dB allocated for connection and splice loss.

[d] 1dB allocated for connection and splice loss.

Table 2-20. *Fiber optic cabling (channel) characteristics at 850nm.*

A channel may contain additional connectors if the optical characteristics of the channel, such as attenuation, modal dispersion, reflections and losses of all connectors and splices meet the specifications. Insertion loss measurements of installed fiber cables are made in compliance with IEC 61280-4-1.

2.6.7 Characteristics of the fiber optic cabling (channel)

The fiber contained within the 100GBASE–SR10 fiber optic cabling shall comply with the specifications and parameters of **Table 2-21**. Many multimode cable types can satisfy these requirements, which provide the resulting channel that also meets the specifications of **Table 2-20**. The fiber optic cabling consists of one or more segments of fiber optic cable and any intermediate connection required to connect the segments together.

Description	OM3 [a]	OM4 [b]	Unit
Nominal core diameter	50		μm
Nominal fiber specification wavelength	850		nm
Effective modal bandwidth (min) [c]	2000	4700	MHz•km
Cabled optical fiber attenuation (max)	3.5		dB/km
Zero dispersion wavelength (λ_0)	$1295 \leq \lambda_0 \leq 1340$		nm
Chromatic dispersion slope (max) (S_0)	0.105 for $1295 \leq \lambda_0 \leq 1310$ and $0.000375 \times (1590 - \lambda_0)$ for $1310 \leq \lambda_0 \leq 1340$		ps/nm^2 km

[a] IEC 60793-2-10 type A1a.2.

[b] IEC 60793-2-10 type A1a.3.

[c] When measured with the launch conditions specified in **Table 2-11**.

Table 2-21. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 2-23**, consists of a mated pair of optical connectors.

The maximum operating distances are based on the allocation of 1.5dB or 1dB total connection and the splice loss. For example, these allocations support two connections, each one has an insertion loss of 0.75dB or 0.5dB, respectively. Connections with lower loss characteristics may be used provided the requirements of **Table 2-20** are met. However, the loss of a single connection shall not exceed 0.75dB.

The maximum discrete reflectance shall be less than -20dB.

2.6.7.1 Medium Dependent Interface (MDI)

The 100GBASE–SR10 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in **Figure 2-23**). The 100GBASE–SR10 PMD is coupled to the fiber optic cabling through one or two connector plugs into the MDI optical connector(s), depending on choice of implementation, as shown in **Figure 2-24**. Example constructions of the MDI include the following:

a. PMD with a connectorized fiber pigtail plugged into an adapter,
b. PMD connector.

2.6.7.2 Optical lane assignments for 100GBASE-SR10

The pinout of the ten transmit and ten receive optical lanes of 100GBASE-SR10 is shown in **Figure 2-24** when looking into the MDI optical connector(s) with the connector keyway feature(s) on top. The single-connector Option A is recommended, the two-connectors Option B and Option C are alternatives. The interface contains 20 active lanes within up to 24 total positions arranged in two rows of 10 or 12 positions. One row is dedicated for the transmit optical lanes and the other row for the receive optical lanes. For the shown 12-position rows, the optical lanes occupy the center ten positions of each row with the outermost positions unused.

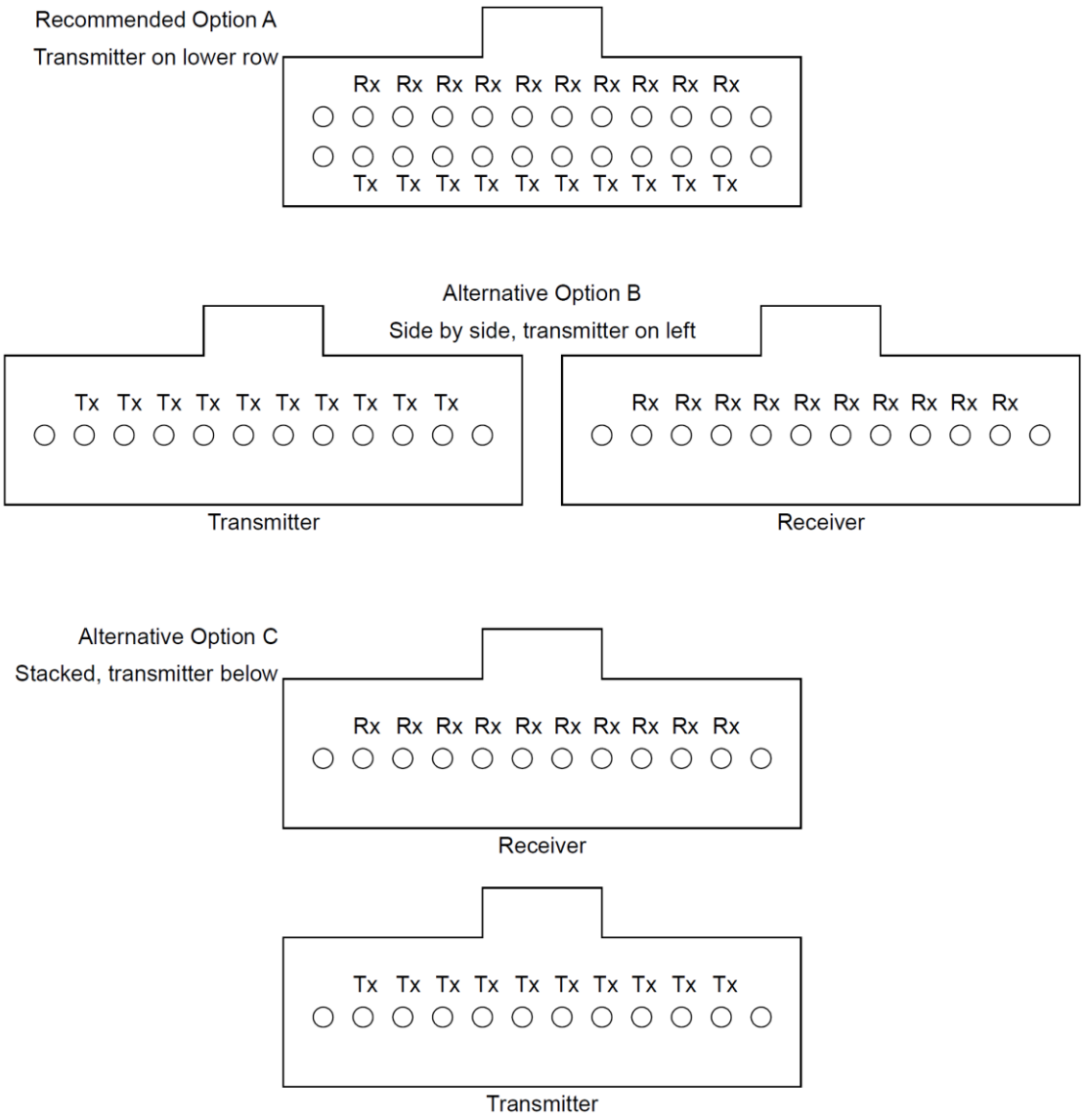


Figure 2-24. *100GBASE-SR10 optical lane assignments*

2.6.7.3 Medium Dependent Interface (MDI) requirements

The MDI adapter or connector shall meet the mechanical specifications of IEC 61754-7 interface 7-3, referred as MPO adapter interface. The plug terminating the optical fiber cabling shall meet the mechanical specifications of IEC 61754-7 interface 7-4, MPO female plug connector flat interface. The MDI shall be mated with the plug on the optical fiber cabling. **Figure 2-25** shows an MPO female plug connector with flat interface, and an MDI as a PMD connector using an MPO adapter interface.

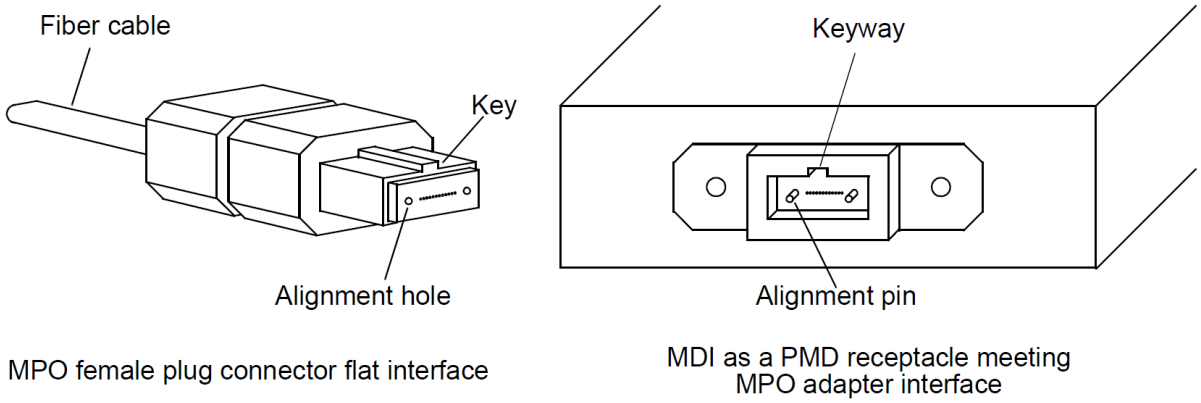


Figure 2-25. *MPO female plug and MDI as a PMD connector using MPO adapter.*

The MDI shall comply with the interface performance specifications of IEC 61753-1-1 and IEC 61753-022-2.

2.7 Physical Medium Dependent (PMD) sublayer and medium, type 100GBASE-LR4 and 100GBASE-ER4

This section specifies the 100GBASE-LR4 PMD and the 100GBASE-ER4 PMD together with the single-mode fiber medium. When forming a complete PHY, a PMD shall be connected to the appropriate PMA as shown in **Table 2-22**, to the medium through the MDI and to the management functions that are optionally accessible through the management interface, or equivalent.

Associated section	100GBASE-LR4	100GBASE-ER4
2.3 - RS	Required	Required
2.3 – CGMII [a]	Optional	Optional
2.4 - PCS for 100GBASE-R	Required	Required
2.5 - PMA for 100GBASE-R	Required	Required
Annex H - CAUI-10	Optional	Optional
Annex I - Chip to module CAUI-10	Optional	Optional
Annex K - CAUI-4	Optional	Optional
Annex L - Chip-to-module CAUI-4	Optional	Optional
Annex F - Energy Efficient Ethernet	Optional	Optional
[a] The CGMII is an optional interface.		

Table 2-22. *PHY sections associated with the 100GBASE-LR4 and 100GBASE-ER4 PMDs.*

Figure 2-26 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC OSI reference model. 100 Gb/s Ethernet is introduced in Section **2.2** and the purpose of each PHY sublayer is summarized in Section **2.3**.

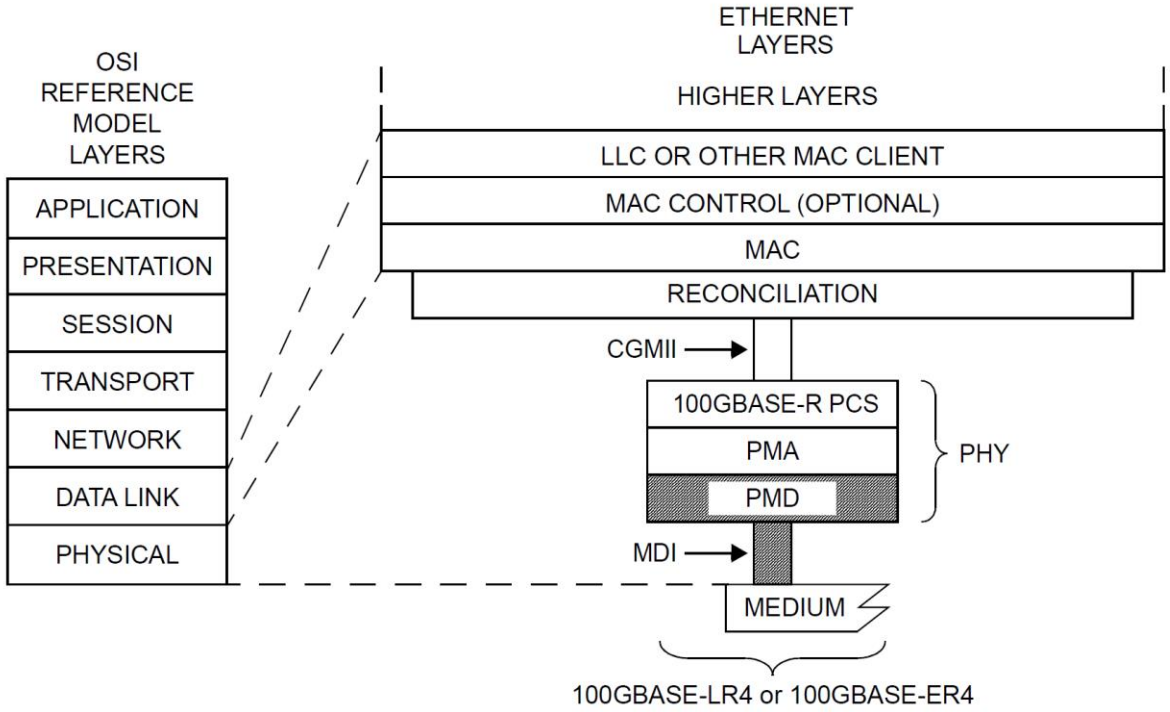


Figure 2-26. *100GBASE-LR4 and 100GBASE-ER4 PMDs relationship to the ISO/IEC OSI reference model and IEEE 802.3 Ethernet model.*

100GBASE-LR4 and 100GBASE-ER4 PHYs with the optional Energy Efficient Ethernet (EEE) fast wake capability can enter the LPI mode to conserve energy during periods of low link utilization (see **Annex F**).

This subsection specifies the services provided by the 100GBASE-LR4 and 100GBASE-ER4 PMDs. The PMD service interface must support the exchange of encoded data between the PMA entity that is allocated just above the PMD, and the PMD entity. The PMD must translate the encoded data to and from signals suitable for the specified medium.

The 100GBASE-LR4 and 100GBASE-ER4 PMDs have four parallel bit streams, from i = 0 to 3.

In the transmit direction, the PMA continuously sends four parallel bit streams to the PMD, one per lane, at a nominal signaling rate of 25.78125GBd. The PMD then converts these streams of bits into the appropriate signals on the MDI.

In the receive direction, the PMD continuously sends four parallel bit streams to the PMA corresponding to the signals received from the MDI, one per lane, at a nominal signaling rate of 25.78125GBd.

The *SIGNAL_DETECT* parameter maps to the *SIGNAL_OK* parameter in the *PMD:IS_SIGNAL.indication(SIGNAL_OK)* inter-sublayer service interface primitive.

The *SIGNAL_DETECT* parameter receives one of the two values, *OK* or *FAIL*. When *SIGNAL_DETECT = FAIL*, then the *rx_bit* parameters are undefined.

NOTE: *SIGNAL_DETECT = OK* does not guarantee that the *rx_bit* parameters are considered as good. It is possible a poor-quality link to provide sufficient light for a *SIGNAL_DETECT = OK* indication and the link still not meet the 10^{-12} BER objective.

The sum of the transmit and receive delays at the one end of the link contributed by the 100GBASE-LR4 or 100GBASE-ER4 PMD, including 2m of fiber in one direction, shall be maximum of 2048-bit times (4 *pause_quanta* or 20.48ns).

The Skew (relative delay) between the lanes shall be kept within the limits, so that the information on the lanes can be reassembled by the PCS. The Skew Variation must also be limited to ensure that a given PCS lane is always traversing the same physical lane. Skew and Skew Variation specified at the points SP1 to SP6 are shown in **Figure 2-5** and **Figure 2-6**.

If the PMD service interface is physically instantiated, so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is also limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

For more information on Skew and Skew Variation see Section **2.2.2**. The tests of Skew and Skew Variation are defined in Section **2.2.2** with the exception that the clock and the data recovery units' high-frequency corner bandwidths are 10MHz.

2.7.1 PMD functional specifications

The 100GBASE-LR4 and 100GBASE-ER4 PMDs perform the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

The PMD block diagram is shown in **Figure 2-27**. The optical transmit signal is defined at the output end of the single-mode fiber optic patch cord (TP2), between 2m and 5m in length. Unless otherwise specified, all transmitter measurements and tests defined in Section **2.7.4** are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see Section **2.7.7.1**). Unless otherwise specified, all receiver measurements and tests defined in Section **2.7.4** are made at TP3.

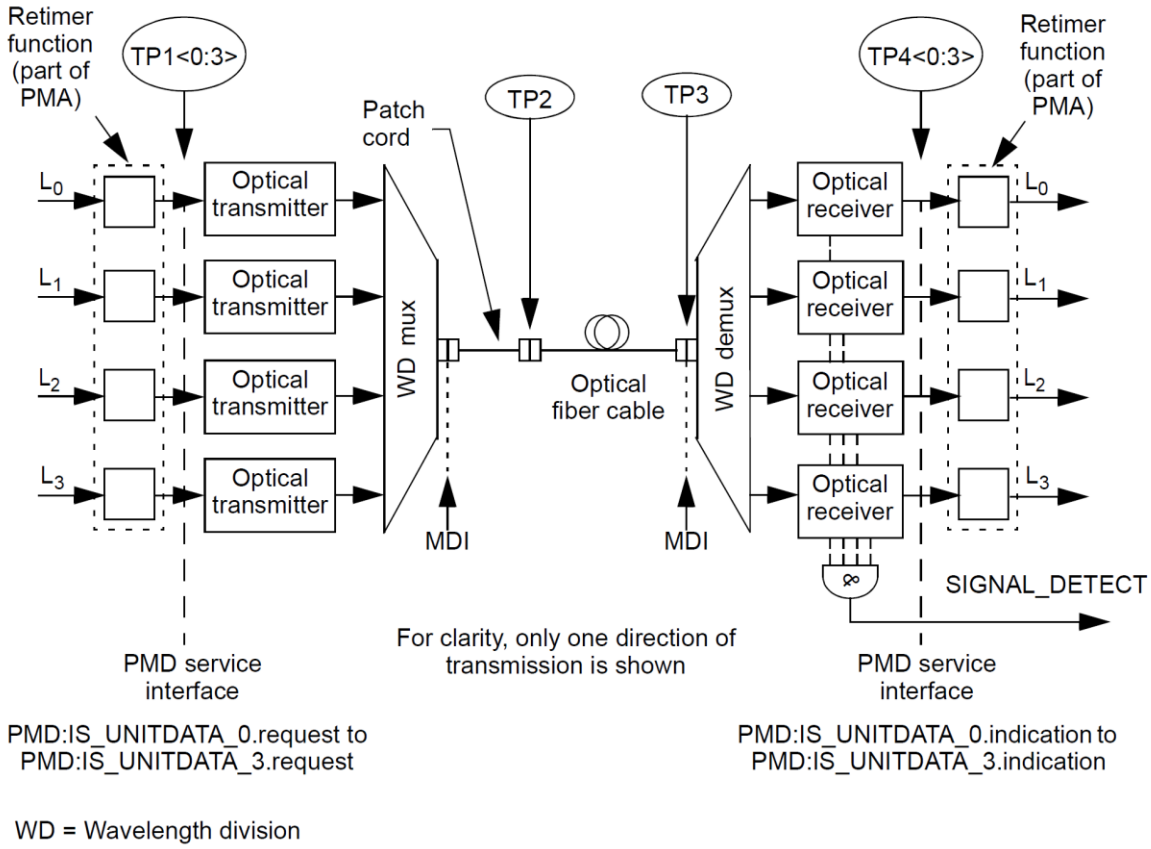


Figure 2-27. *Block diagram for 100GBASE-LR4 and 100GBASE-ER4 transmit/receive paths.*

2.7.1.1 PMD transmit function

The PMD Transmit function shall convert the four bit streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_3.request* into four separate optical signal streams. The four optical signal streams shall then be wavelength division multiplexed and delivered to the MDI, all according to the transmit optical specifications defined in this Section. The higher optical power level in each signal stream shall correspond to *tx_bit* = *one*.

2.7.1.2 PMD receive function

The PMD Receive function shall demultiplex the composite optical signal stream received from the MDI into four separate optical signal streams. The four optical signal streams shall then be converted into four bit streams for delivery to the PMD service interface using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_3.indication*, all according to the receive optical specifications defined in this Section. The higher optical power level in each signal stream shall correspond to *rx_bit* = *one*.

2.7.1.3 PMD global signal detect function

The PMD global signal detect function reports the state of *SIGNAL_DETECT* parameter via the PMD service interface. The *SIGNAL_DETECT* parameter is signaled continuously, while the *PMD:IS_SIGNAL.indication* message is generated when a change occurs in the value of *SIGNAL_DETECT* parameter. The *SIGNAL_DETECT* parameter maps to the *SIGNAL_OK* parameter in the inter-sublayer service interface primitives.

The *SIGNAL_DETECT* is a global indicator, which define the presence of optical signals on all four lanes. The value of the *SIGNAL_DETECT* parameter is generated according to the conditions defined in **Table 2-23**. The PMD receiver is not required to verify whether or not a compliant 100GBASE-R signal is being received.

Receive conditions	SIGNAL_DETECT value
For any lane; Average optical power at TP3 ≤ -30dBm	FAIL
For all lanes, [(Optical power at TP3 ≥ receiver sensitivity (max) in OMA in **Table 2-27**)	OK

AND (Compliant 100GBASE–R signal input)]	
All other conditions	Unspecified

Table 2-23. *SIGNAL_DETECT value definition.*

When the *SIGNAL_DETECT* parameter is setting, implementations must provide adequate margin between the input optical power level, where the *SIGNAL_DETECT* parameter is OK, and the inherent noise level of the PMD, including the effects of crosstalk, power supply noise, etc.

Various implementations of the *SIGNAL_DETECT* function are permitted, including implementations that generate the *SIGNAL_DETECT* parameter values in response to the amplitude of the modulation of the optical signal and implementations that respond to the average optical power of the modulated optical signal.

2.7.1.4 PMD lane-by-lane signal detect function

Various implementations of the *SIGNAL_DETECT* function are permitted. When the MDIO is implemented, each *PMD_signal_detect_i*, where *i* represents the lane number in the range *0 to 3*, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of **Table 2-23**.

2.7.2 Wavelength-division-multiplexed lane assignments

The wavelength range for each lane of the 100GBASE-LR4 and 100GBASE-ER4 PMDs is defined in **Table 2-24**. The center frequencies are members of the frequency grid for 100GHz spacing and above defined in ITU-T G.694.1 and are spaced at 800GHz.

Lane	Center frequency	Center wavelength	Wavelength range
L_0	231.4THz	1295.56nm	1294.53 to 1296.59nm
L_1	230.6THz	1300.05nm	1299.02 to 1301.09nm
L_2	229.8THz	1304.58nm	1303.54 to 1305.63nm
L_3	229THz	1309.14nm	1308.09 to 1310.19nm

Table 2-24. *Wavelength-division-multiplexed lane assignments.*

2.7.3 PMD to MDI optical specifications for 100GBASE-LR4 and 100GBASE-ER4

The operating ranges for the 100GBASE-LR4 and 100GBASE-ER4 PMDs are defined in **Table 2-25**. A 100GBASE-LR4 or 100GBASE-ER4 compliant PMD operates on type B1.1, B1.3, or B6 a single-mode fibers according to the specifications defined in **Table 2-33**. A PMD that exceeds the operating range requirement while

meeting all other optical specifications is considered compliant (e.g., a 100GBASE-LR4 PMD operating at 12.5km meets the operating range requirement of 2m to 10km). The 100GBASE-ER4 PMD interoperates with the 100GBASE-LR4 PMD provided that the channel requirements for 100GBASE-LR4 are met.

PMD type	Required operating range
100GBASE-LR4	2m to 10km
100GBASE-ER4	2m to 30km
	2m to 40km [a]
a Links longer than 30km for the same link power budget are considered engineered links. Attenuation for such links needs to be less than the worst case specified for B1.1, B1.3, or B6 a single-mode fiber.	

Table 2-25. *100GBASE-LR4 and 100GBASE-ER4 operating ranges.*

2.7.3.1 100GBASE-LR4 and 100GBASE-ER4 transmitter optical specifications

The 100GBASE-LR4 and 100GBASE-ER4 transmitter shall meet the specifications defined in **Table 2-26** as per the definitions in Section **2.7.4**.

Description	100GBASE-LR4	100GBASE-ER4	Unit
Signaling rate, each lane (range)	25.78125 ± 100ppm		GBd
Lane wavelengths (range)	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19		nm
Side-mode suppression ratio (SMSR), (min)	30		dB
Total average launch power (max)	10.5	8.9	dBm
Average launch power, each lane (max)	4.5	2.9	dBm
Average launch power, each lane [a] (min)	-4.3	-2.9	dBm
Optical Modulation Amplitude (OMA), each lane (max)	4.5		dBm
Optical Modulation Amplitude (OMA), each lane (min)	-1.3 [b]	0.1	dBm
Difference in launch power between any two lanes (OMA) (max)	5	-	dB
Difference in launch power between any two lanes (Average and OMA) (max)	-	3.6	dB

Launch power in OMA minus TDP, each lane (min)	-2.3	-	dBm
Transmitter and dispersion penalty (TDP), each lane (max)	2.2	2.5	dB
Average launch power of OFF transmitter, each lane (max)	-30		dBm
Extinction ratio (min)	4	8	dB
$RIN_{20}OMA$ (max)	-130		dB/Hz
Optical return loss tolerance (max)	20		dB
Transmitter reflectance [c] (max)	-12		dB
Transmitter eye mask definition {X1, X2, X3, Y1, Y2, Y3}	{0.25, 0.4, 0.45, 0.25, 0.28, 0.4}		
[a] Average launch power, each lane (min) is informative and it is not the principal indicator of signal strength. [b] The OMA (min) must exceed this value, even if the TDP < 1dB. [c] Transmitter reflectance is defined at the transmitter side.			

Table 2-26. *100GBASE-LR4 and 100GBASE-ER4 transmit characteristics.*

2.7.3.2 100GBASE-LR4 and 100GBASE-ER4 receive optical specifications

The 100GBASE-LR4 receiver shall meet the specifications defined in **Table 2-27** per the definitions in Section **2.7.4**. The 100GBASE-ER4 receiver shall meet the specifications defined in **Table 2-27** per the definitions in Section **2.7.4**.

Description	100GBASE-LR4	100GBASE-ER4	Unit
Signaling rate, each lane (range)	25.78125 ± 100ppm		GBd
Lane wavelengths (range)	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19		nm
Damage threshold [a]	5.5		dBm
Average receive power, each lane (max)	4.5 [b]		dBm
Average receive power, each lane [c] (min)	-10.6	-2.09	dBm
Receive power, each lane (OMA) (max)	4.5		dBm
Difference in receive power between any two lanes (OMA) (max)	5.5	-	dB
Difference in receive power between any two lanes (Average and OMA) (max)	-	4.5	dB

Receiver reflectance (max)	-26		dB
Receiver sensitivity (OMA), each lane [d] (max)	-8.6	-21.4	dBm
Receiver 3 dB electrical upper cutoff frequency, each lane (max)	31		GHz
Stressed receiver sensitivity (OMA), each lane [e] (max)	-6.8	-17.9	dBm
Conditions of stressed receiver sensitivity test			
Vertical eye closure penalty, [f] each lane	1.8	3.5	dB
Stressed eye J2 Jitter, [f] each lane	0.3		UI
Stressed eye J9 Jitter, [f] each lane	0.47		UI

[a] The receiver can tolerate, without damage, continuous exposure to an optical input signal, which has this average power level.

[b] The average receive power, each lane (max) for 100GBASE-ER4 is larger than the 100GBASE-ER4 transmitter value allowing compatibility with 100GBASE-LR4 units at short distances.

[c] Average receive power, each lane (min) is informative, and it is not the principal indicator of signal strength.

[d] Receiver sensitivity (OMA), each lane (max) is informative only.

[e] Measured with conformance test signal at TP3 (see Section **2.7.4.13**) for BER = 10^{-12}.

[f] Vertical eye closure penalty, stressed eye J2 Jitter, and stressed eye J9 Jitter are test conditions to measure stressed receiver sensitivity. They are not characteristics of the receiver.

Table 2-27. *100GBASE-LR4 and 100GBASE-ER4 receive characteristics.*

2.7.3.3 100GBASE-LR4 and 100GBASE-ER4 illustrative link power budgets

Illustrative link power budgets and penalties for 100GBASE-LR4 and 100GBASE-ER4 channels are shown in **Table 2-28**.

Parameter	100GBASE-LR4	100GBASE-ER4		Unit
Power budget (for maximum TDP)	8.5	-		dB
Power budget	-	21.5		dB
Operating distance	10	30	40 [a]	km
Channel insertion loss	6.3 [b]	15	18	dB
Maximum discrete reflectance	-26	-26		dB
Allocation for penalties [c] (for maximum TDP)	2.2	-		dB

Allocation for penalties [c]	-	3.5		dB
Additional insertion loss allowed	0	3	0	dB

[a] Links longer than 30km are considered as engineered links. Attenuation for such links needs to be less than the worst case for B1.1, B1.3, or B6 single mode cabled optical fiber.

[b] The channel insertion loss is calculated using the maximum distance specified in **Table 2-25** for 100GBASE-LR4 and fiber attenuation of 0.43dB/km at 1295nm plus an allocation for connection and splice loss given in Section **2.7.7**.

[c] The link penalties are used for link budget calculations.

Table 2-28. *100GBASE-LR4 and 100GBASE-ER4 illustrative link power budgets.*

2.7.4 Definition of optical parameters and measurement methods

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified.

Compliance shall be achieved in normal operation, when specific test patterns are defined for measurement consistency and enable measurement of some parameters. **Table 2-29** gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subsections in which each parameter is defined. Any of the test patterns given for a particular test in **Table 2-29** may be used to perform that test. The test patterns used in this Section are shown in **Table 2-30**.

Parameter	Pattern	Related Section
Wavelength	3, 5 or valid 100GBASE-R signal	**2.7.4**
Side mode suppression ratio	3, 5 or valid 100GBASE-R signal	**2.7.4.1**
Average optical power	3, 5 or valid 100GBASE-R signal	**2.7.4.2**
Optical modulation amplitude (OMA)	Square wave or 4	**2.7.4.3**
Transmitter and dispersion penalty (TDP)	3 or 5	**2.7.4.4**
Extinction ratio	3, 5 or valid 100GBASE-R signal	**2.7.4.9**
$RIN_{20}OMA$	Square wave or 4	**2.7.4.10**
Transmitter optical waveform	3, 5 or valid 100GBASE-R signal	**2.7.4.11**
Stressed receiver sensitivity	3 or 5	**2.7.4.13**

Calibration of OMA for receiver tests	Square wave or 4	**2.7.4.14**
Vertical eye closure penalty calibration	3 or 5	**2.7.4.14**
Receiver 3 dB electrical upper cutoff frequency	3, 5 or valid 100GBASE-R signal	**2.7.4.14**

Table 2-29. *Test-pattern definitions and related sections.*

Pattern	Pattern description
Square wave	Square wave (8 ones, 8 zeros)
3	PRBS31
4	PRBS9
5	Scrambled idle

Table 2-30. *Test patterns.*

2.7.4.1 Wavelength and side mode suppression ratio (SMSR)

The wavelength and SMSR of each optical lane shall be within the ranges given in **Table 2-24** if measured as per IEC 61280-1-3. The lane under test is modulated using the test pattern defined in **Table 2-29**.

2.7.4.2 Average optical power

The average optical power of each lane shall be within the limits given in **Table 2-26** if measured using the methods given in IEC 61280-1-1, with the sum of the optical power from all of the lanes not under test below –30dBm, per the test setup in **Figure 2-28**.

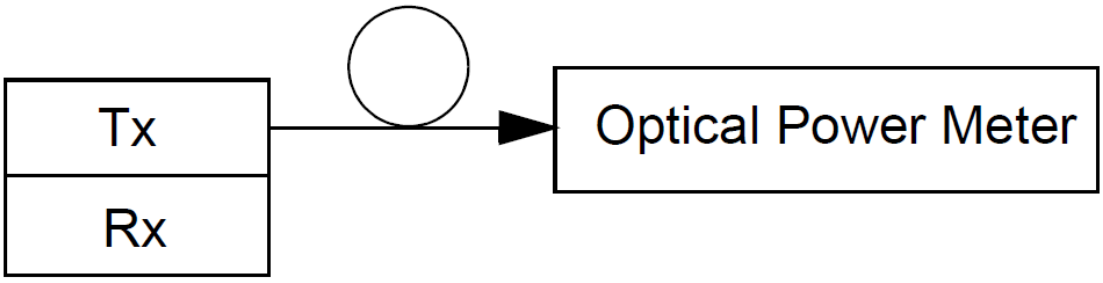


Figure 2-28. *Optical power measurement test set-up.*

2.7.4.3 Optical Modulation Amplitude (OMA)

OMA shall be as defined in Section **2.6.4.3** for measurement with a square wave (8 ones, 8 zeros) test pattern or from the variable *MeasuredOMA* for measurement with a PRBS9 test pattern, with the exception that each lane can be tested individually with

the sum of the optical power from all of the lanes not under test being below –30dBm, or in case other lanes are operating, a suitable optical filter can be used to separate the lane under test.

2.7.4.4 Transmitter and dispersion penalty (TDP)

TDP shall be as defined in Section **2.6.4.4** with the exception that each optical lane is tested individually using an optical filter to separate the lane under test from the others. The measurement procedure for 100GBASE-LR4 and 100GBASE-ER4 is detailed in Section **2.7.4.5** to Section **2.7.4.8**.

The optical filter passband ripple shall be limited to 0.5dB peak-to-peak and the chosen isolation must be such that the ratio of the power in the measured lane to the sum of the powers of all of the other lanes will be greater than 20dB (see ITU-T G.959.1 **Annex B**). The lanes which are not under test shall be operating with PRBS31 or valid 100GBASE-R bit streams.

2.7.4.5 Reference transmitter requirements

The reference transmitter shall be a high-quality instrument-grade device, which can be implemented by using a CW laser modulated by a high-performance modulator. The basic requirements are as follows:

a. Rise/fall times of less than 12ps at 20% to 80%.
b. The output optical eye is symmetric and passes the transmitter optical waveform test of Section **2.7.4.11**.
c. In the center 20% region of the eye, the worst-case vertical eye closure penalty as defined in Section **2.6.4.9** is less than 0.5dB.
d. Total Jitter less than 0.2 UI peak-to-peak.
e. RIN of less than -138dB/Hz.

2.7.4.6 Channel requirements

The transmitter is tested using an optical channel that meets the requirements listed in **Table 2-31**.

PMD type	Dispersion [a] (ps/nm)		Insertion Loss [b]	Optical return loss [c]	Max mean DGD
	Minimum	Maximum			
100GBASE-LR4	$0.2325 \cdot \lambda \cdot [1-(1324/\lambda)^4]$	$0.2325 \cdot \lambda \cdot [1-(1300/\lambda)^4]$	Minimum	20dB	0.8ps
100GBASE-ER4	$0.93 \cdot \lambda \cdot [1-(1324/\lambda)^4]$	$0.93 \cdot \lambda \cdot [1-(1300/\lambda)^4]$	Minimum	20dB	0.8ps

[a] The dispersion is measured for the wavelength of the device under test (λ is in nm). The coefficient assumes 10km for 100GBASE-LR4 and 40km for 100GBASE-ER4.

[b] Assume that there is nothing to stress the sensitivity of the BERT's optical receiver.

[c] The optical return loss takes place at TP2.

Table 2-31. *Transmitter compliance channel specifications.*

A 100GBASE-LR4 or 100GBASE-ER4 transmitter must be compliant with a total dispersion, with values at least as negative as the ones of "minimum dispersion" and at least as positive as the ones of the "maximum dispersion" columns specified in **Table 2-31** for the wavelength of the device under test. This can be achieved with channels consisting of fibers with lengths chosen to meet the dispersion requirements.

To verify that the fiber complies with the correct dispersion requirements, the measurement method must be according to the IEC 60793-1-42. The measurement must take place in the linear power regime of the fiber.

The channel provides an optical return loss specified in **Table 2-31**. The state of polarization of the back reflection must be adjusted to create the greatest RIN.

The mean Differential Group Delay (DGD) of the channel is to be less than the value specified in **Table 2-31**.

2.7.4.7 Reference receiver requirements

The reference receiver is required to have the bandwidth given in Section **2.7.4.11**. The sensitivity of the reference receiver is limited by Gaussian noise. The receiver shall have the minimum threshold offset, deterministic jitter, baseline wander, dead band, hysteresis or other distortions. Decision sampling has minimal uncertainty and setup/hold times.

The nominal sensitivity of the reference receiver, *S*, is measured in OMA using the setup of **Figure 2-20** without the test fiber and with the transversal filter removed. The sensitivity *S* shall be corrected for any reference transmitter impairments, including any vertical eye closure. It is measured when it is sampling at the eye center or corrected for the off-center sampling and it is calibrated at the wavelength of the transmitter under test.

The determination of the center of the eye is required for all transmitter and dispersion penalty measurements. The center of the eye is defined as the halfway time between the left and the right sampling points within the eye, where the measured BER is greater than or equal to 1×10^{-3}.

The Clock Recovery Unit (CRU) used in the TDP measurement shall have a corner frequency of 10MHz and a slope of 20dB/decade. When using a CRU as a clock for the measurement of BER, passing of the low frequency jitter from the data to the clock removes this low-frequency jitter from the measurement.

2.7.4.8 Test procedure

The test procedure is as defined in Section **2.6.4.4** with the exception that all lanes are operational in both directions (transmit and receive), each lane is tested individually using an optical filter to separate the lane under test from the others and the BER of 1×10^{-12} is for the lane under test on its own.

2.7.4.9 Extinction ratio

The extinction ratio of each lane shall be within the limits given in **Table 2-26** if measured using the methods specified in IEC 61280-2-2, with the sum of the optical power from all of the lanes not under test below -30dBm. The extinction ratio is measured using the test pattern defined in **Table 2-29**.

NOTE: Extinction ratio and OMA are defined with different test patterns (see **Table 2-29**).

2.7.4.10 Relative Intensity Noise ($RIN_{20}OMA$)

RIN shall be as defined by the measurement methodology of Section **2.6.4.5** with the following exceptions:

a. The optical return loss shall be 20dB.
b. Each lane shall be tested individually with the sum of the optical power from all of the lanes not under test being below -30dBm, or if other lanes are operating, where a suitable optical filter can be used to separate the lanes under test.
c. The upper -3dB limit of the measurement apparatus shall be approximately equal to the signaling rate (i.e., 25.8GHz).

2.7.4.11 Transmitter optical waveform (transmit eye)

The required shape characteristics of the optical transmitter pulse are specified in the form of a mask of the transmitter eye diagram as shown in **Figure 2-22**. The transmitter optical waveform of a port transmitting the test pattern specified in **Table 2-17** shall meet specifications according to the methods specified in Section **2.6.4.7** with the exception that the clock recovery unit's high-frequency corner bandwidth is 10MHz. The filter nominal reference frequency f_r is 19.34GHz and the filter tolerances are as specified for STM-64 in ITU-T G.691. Compensation can be made for the

variation of the reference receiver filter response from an ideal fourth-order Bessel-Thomson response.

2.7.4.12 Receiver sensitivity

The receiver sensitivity is only informative, and compliance is not required. If measured, the test signal should have negligible impairments such as ISI, rise/fall times, jitter, and RIN, while the normative requirement for receivers is stressed receiver sensitivity.

2.7.4.13 Stressed receiver sensitivity

The stressed receiver sensitivity shall be within the limits given in **Table 2-27**, if measured using the method defined in Section **2.6.4.9** with the following exceptions:

a. Added sinusoidal jitter is as specified in **Table 2-32**.
b. The stressed eye J2 Jitter, stressed eye J9 Jitter, and vertical eye closure penalty are as given in **Table 2-27**.
c. The test pattern is as given in **Table 2-29**.
d. The reference receiver used to verify the conformance test signal is required to have the bandwidth given in Section **2.7.4.11**.

Frequency range	Sinusoidal jitter, peak-to-peak (UI)
f < 100kHz	Not specified
100kHz < f ≤ 10MHz	$5 \times 10^5/f$
10MHz < f < 10 LB [a]	0.05
[a] LB is the loop bandwidth, upper frequency bound for added sine jitter should be at least 10 times the loop bandwidth of the receiver being tested.	

Table 2-32. *Applied sinusoidal jitter.*

2.7.4.14 Receiver 3 dB electrical upper cutoff frequency

The receiver 3dB electrical upper cutoff frequency shall be within the limits given in **Table 2-27**. Each optical lane is measured using an optical signal or signals with its/their wavelength within the specified wavelength range of the lane to be tested.

2.7.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

100GBASE-LR4 and 100GBASE-ER4 optical transceivers shall conform to Hazard Level 1 laser requirements in compliance with IEC 60825-1 and IEC 60825-2, under any condition of operation.

Laser safety standards and regulations require that the manufacturer of the laser product shall provide information about the product's laser, safety features, labeling, use, maintenance, and service.

It is recommended that proper installation practices, as defined by applicable local codes and regulation, shall be followed in every instance in which such practices are applicable.

It is recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of this Section will be met.

A system integrating either a 100GBASE-LR4 or a 100GBASE-ER4 PMD shall comply with the applicable local codes and regulations for the limitation of electromagnetic interference.

The optical link shall operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration).

2.7.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 2-29**.

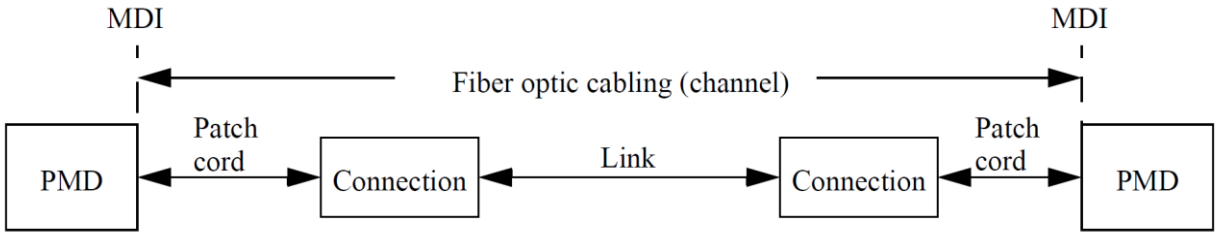


Figure 2-29. *Fiber optic cabling model.*

The channel insertion loss is given in **Table 2-33**. The channel can contain additional connectors, if its optical characteristics, such as the attenuation, the dispersion, the reflections, and the polarization mode dispersion comply with the specifications. Insertion loss measurements of the installed fiber optic cables shall be made in accordance with IEC 61280-1-1. The term "channel" is used for consistency with the generic cabling standards.

Description	100GBASE-LR4	100GBASE-ER4		Unit
Operating distance (max)	10	30	40	km
Channel insertion loss [a, b] (max)	6.3	18	18	dB
Channel insertion loss (min)	0	0		dB
Positive dispersion [b] (max)	9.5	28	36	ps/nm

Negative dispersion [b] (min)	-28.5	-85	-114	ps/nm
DGD_max [c]	8	10.3	10.3	ps
Optical return loss (min)	21	21	21	dB

[a] These values of the channel insertion loss include the cable, the connectors, and the splices.

[b] Over the wavelength range 1294.53nm to 1310.19nm.

[c] Differential Group Delay (DGD) is the time difference at the reception point between the fractions of a pulse, which were transmitted in the two main states of polarization of an optical signal. DGD_max is the maximum DGD that the system shall tolerate.

Table 2-33. *Fiber optic cabling (channel) characteristics.*

2.7.7 Characteristics of the fiber optic cabling (channel)

The 100GBASE-LR4 and 100GBASE-ER4 fiber optic cabling shall meet the specifications defined in **Table 2-33**. The fiber optic cabling shall be comprised of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

The fiber optic cable requirements are satisfied by fibers, which are complying with the IEC 60793-2-50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-mode), or type B6_a (bend insensitive) standards or the requirements in **Table 2-34** where they differ.

Description	Value	Unit
Nominal fiber specification wavelength	1310	nm
Cabled optical fiber attenuation (max)	0.43 [a] or 0.5 [b]	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0,093	ps/nm^2 km

a The attenuation of 0.43dB/km at 1295nm for the optical fiber cables is derived from Appendix I of ITU-T G.695.

b The attenuation of 0.5dB/km is referred to the outdoor cables, which shall be in compliance with CENELEC EN 50173-1. Using 0.5dB/km may not support operation at 10km for 100GBASE-LR4 or 40km for 100GBASE-ER4.

Table 2-34. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 2-29**, consists of a mated pair of optical connectors.

The maximum link distance is based on an allocation of 2dB total connection and splice loss. For example, this channel supports four connections with an average

insertion loss per connection of 0.5dB. Connections with different loss characteristics may be used provided the requirements of **Table 2-33** are met. The value of the maximum discrete reflectance shall be less than -26dB.

2.7.7.1 Medium Dependent Interface (MDI) requirements

The 100GBASE-LR4 or 100GBASE-ER4 PMD is connected to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in **Figure 2-29**). Examples of an MDI include the following:

a. Connectorized fiber pigtail.
b. PMD connector.

When the MDI is a connection comprised of a plug and an adapter, it shall meet the interface performance specifications of IEC 61753-1-1 and IEC 61753-021-2.

NOTE: Transmitter compliance testing is performed at TP2 as defined in Section **2.7.1**, not at the MDI.

2.8 Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs

This Section specifies a Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs.

Figure 2-30 shows the relationship of the RS-FEC sublayer to the ISO/IEC OSI reference model.

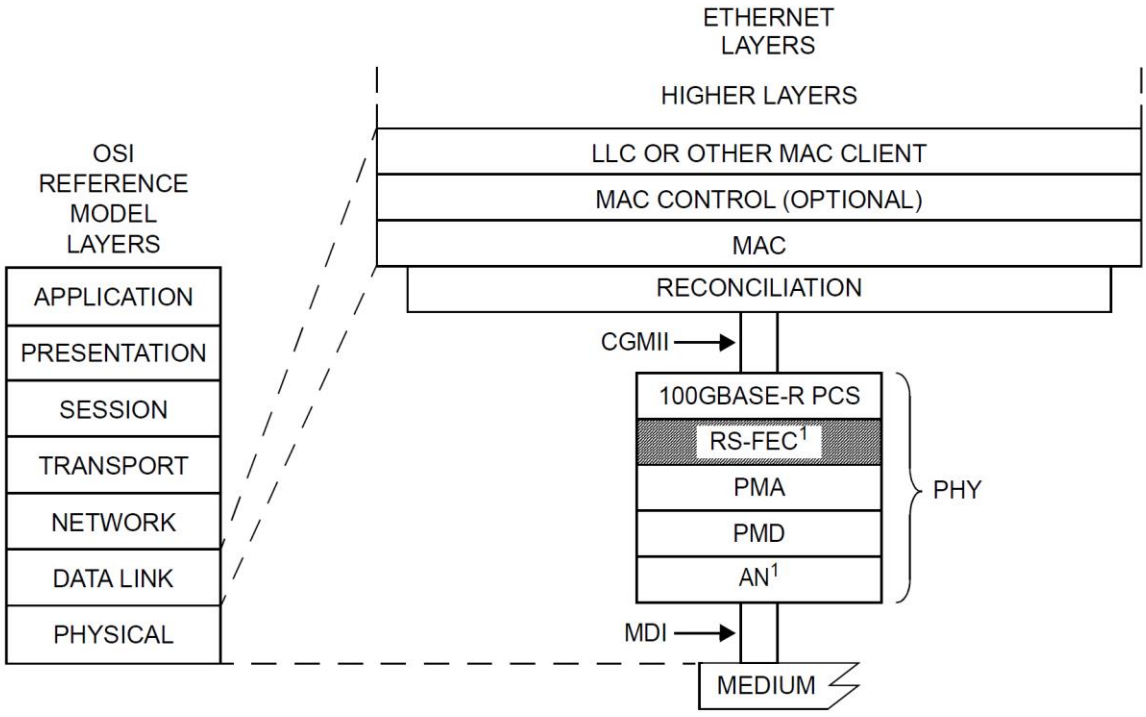


Figure 2-30. *RS-FEC relationship to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model.*

2.8.1 FEC service interface

This Section specifies the services provided by the RS-FEC sublayer. The FEC service interface allows the PCS to transfer information to and from the RS-FEC. The PCS can be connected to the RS-FEC using an optional instantiation of the PMA service interface (refer to **Annex H** or **Annex K**) in which case a PMA is the client of the FEC service interface.

The FEC service interface is provided as an instance of the inter-sublayer service interface. The FEC service interface primitives are the following:

- *FEC:IS_UNITDATA_i.request*
- *FEC:IS_UNITDATA_i.indication*
- *FEC:IS_SIGNAL.indication*

The RS-FEC uses 20 parallel bit streams, hence *i* = 0 to 19. The PCS (or PMA) continuously transmits 20 parallel bit streams to the RS-FEC, one bit stream per lane,

each at the nominal signaling rate of 5.15625GBd. Also, the RS-FEC continuously transmits 20 parallel bit streams to the PCS (or PMA), one bit stream per lane, each at the nominal signaling rate of 5.15625GBd.

The *SIGNAL_OK* parameter of the *FEC:IS_SIGNAL.indication* primitive can receive one of two values: *OK* or *FAIL*. The value is set to *OK,* when the FEC receive function is identified the codeword boundaries as indicated by *fec_align_status* equal to true. That value is set to *FAIL* when the *FEC* receive function is unable to establish codeword boundaries as indicated by *fec_align_status* equal to false. When *SIGNAL_OK* is *FAIL*, then the *rx_bit* parameters of the *FEC:IS_UNITDATA_i.indication* primitives will be undefined.

If the optional EEE deep sleep capability is supported, then the FEC service interface has four additional primitives as follows:

- *FEC:IS_TX_MODE.request*
- *FEC:IS_RX_MODE.request*
- *FEC:IS_RX_TX_MODE.indication*
- *FEC:IS_ENERGY_DETECT.indication*

When the *tx_mode* parameter of the *FEC:IS_TX_MODE.request* primitive is *QUIET* or *ALERT*, the RS-FEC sublayer can disable transmit functional blocks to conserve energy. Otherwise, the RS-FEC transmit function operates normally. The value of *tx_mode* is transferred to the client sublayer via the *PMA:IS_TX_MODE.request* primitive.

When the *rx_mode* parameter of the *FEC:IS_RX_MODE.request* primitive is *QUIET*, the RS-FEC sublayer can disable receive functional blocks to conserve energy. Otherwise, the RS-FEC receive function operates normally. The value of *rx_mode* is transferred to the client sublayer via the *PMA:IS_RX_MODE.request* primitive.

The *rx_tx_mode* parameter of the *FEC:IS_RX_TX_MODE.indication* primitive is used to communicate the link partner's value of *tx_mode* as inferred by the PMA. It is allocated that the value is received via the *PMA:IS_RX_TX_MODE.indication* primitive.

The *energy_detect* parameter of the *FEC:IS_ENERGY_DETECT.indication* primitive is used to inform that the PMD has detected the return of energy on the

interface after a period of quiescence. It is allocated that the value is received via the *PMA:IS_ENERGY_DETECT.indication* primitive.

2.8.2 PMA compatibility

The RS-FEC sublayer requires that the PMA service interface consist of exactly four upstream lanes and exactly four downstream lanes. Hence, the RS-FEC sublayer can be a client of the PMA sublayer defined in Section **2.5**, when the PMA service interface width, *p*, is set to 4.

In addition, all PMA service interfaces between the RS-FEC sublayer and the PMD sublayer are required to consist of four or fewer upstream lanes and four or fewer downstream lanes. A consequence of this constraint is that a physical instantiation of the ten-lane PMA service interface (CAUI-10) may not be used below the RS-FEC sublayer.

2.8.3 Delay constraints

The maximum delay contributed by the RS-FEC sublayer (sum of transmit and receive delays at one end of the link) shall be no more than 40960bit times (80 *pause_quanta* or 409.6ns). A description of overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **2.2.1** and its references.

2.8.4 Functions within the RS-FEC sublayer

A functional block diagram of the RS-FEC sublayer is shown in **Figure 2-31**.

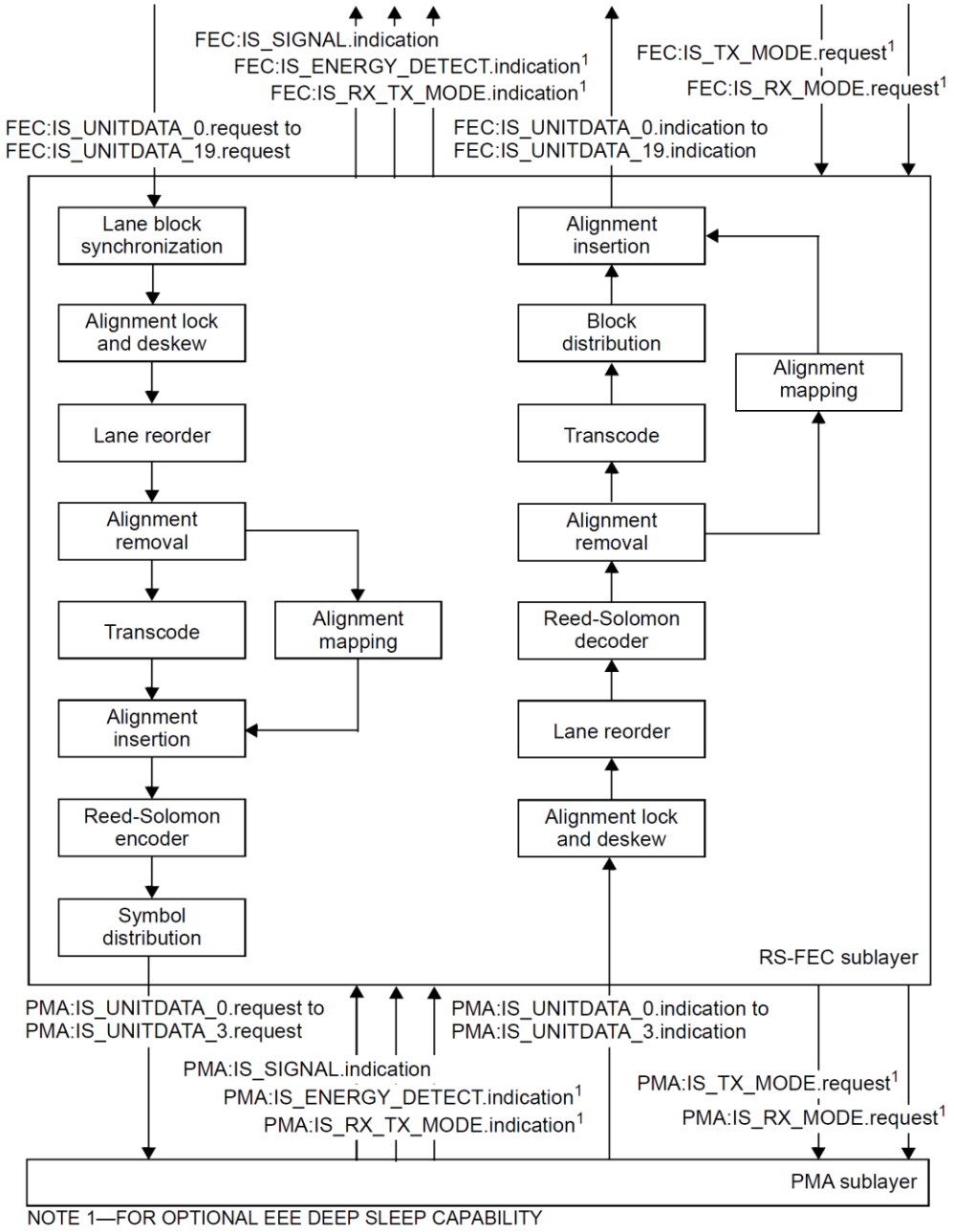


Figure 2-31. *Functional block diagram.*

2.8.4.1 Transmit function

The RS-FEC transmit function consists of 20 bit streams by concatenating the bits from each of the 20 *FEC:IS_UNITDATA_i.request* primitives in the order they are received. It locks the 66-bit blocks in each bit stream by using the sync headers and outputs 66-bit blocks. Block lock is obtained as specified in the block lock state diagram shown in **Figure 2-32**.

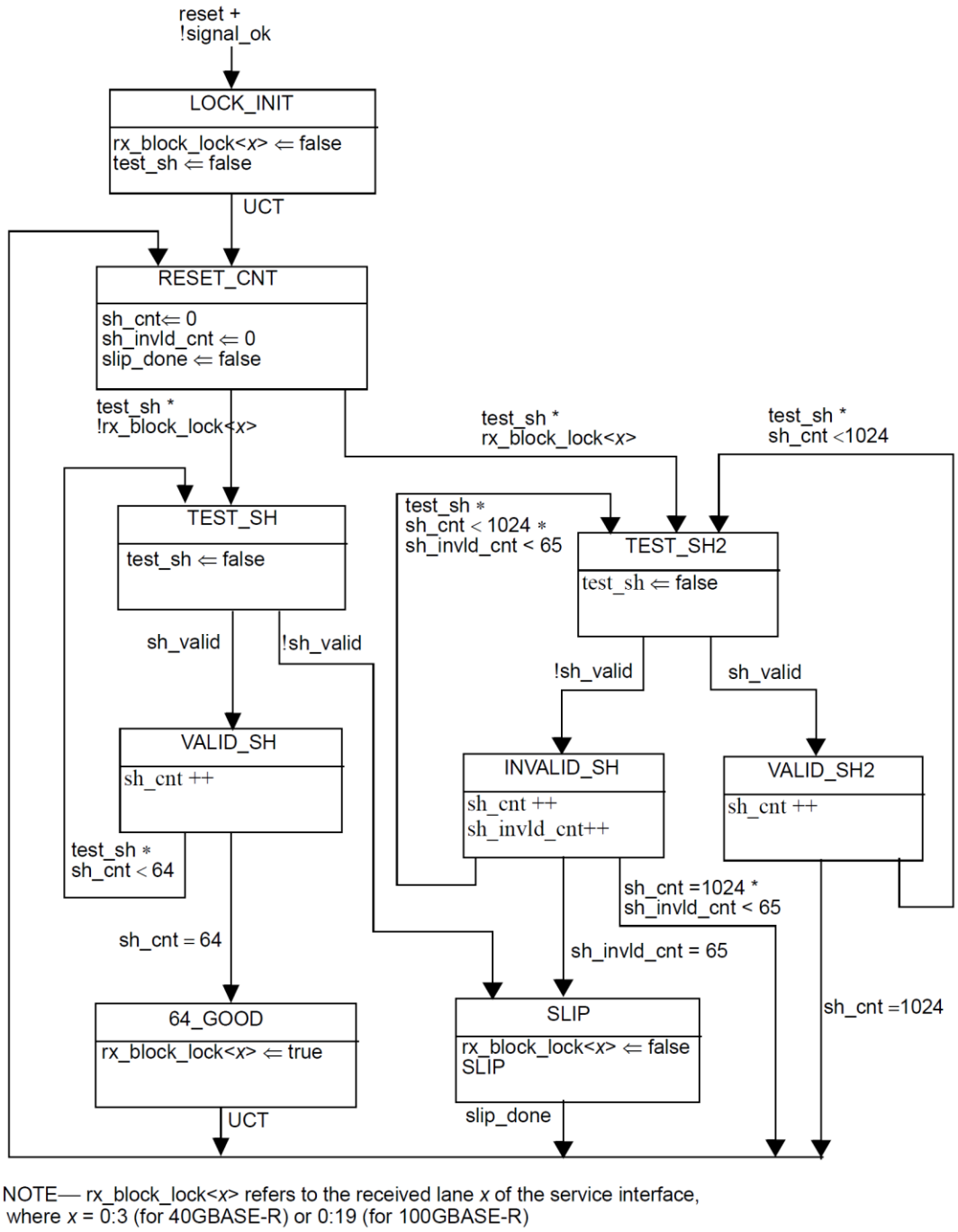
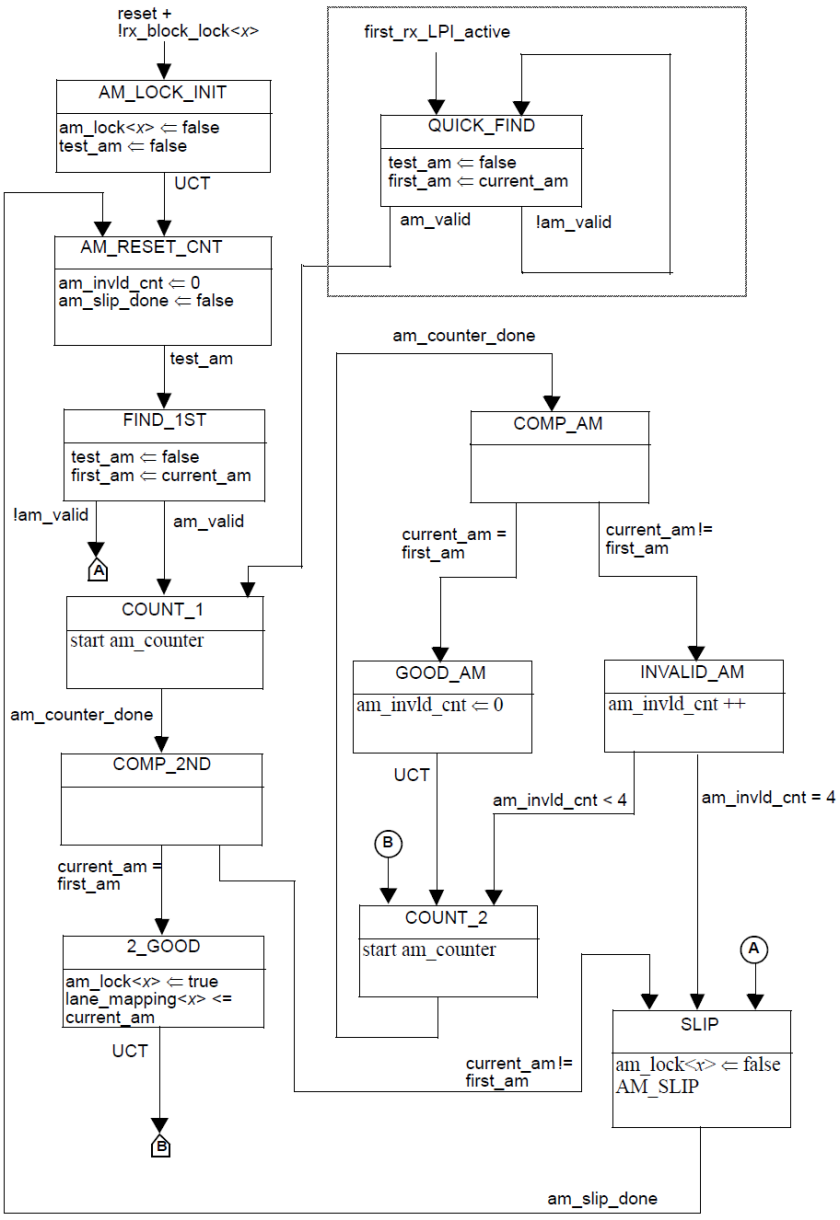


Figure 2-32. *Block lock state diagram*

Once the RS-FEC transmit function achieves block lock on a PCS lane, it then begins obtaining alignment marker lock as specified by the alignment marker lock state diagram shown in **Figure 2-33**. This process identifies the number of the PCS lane received on a particular lane of the service interface. After alignment marker lock is achieved on all 20 lanes, all inter-lane Skew is removed as specified by the PCS deskew state diagram shown in **Figure 2-34**. The RS-FEC transmit function supports

a maximum Skew of 49ns between the PCS lanes and the maximum Skew Variation of 400ps. Skew and Skew Variation are defined in Section **2.2.2**.



NOTE 1— am_lock<*x*> refers to the received lane *x* of the service interface, where *x* = 0:3 (for 40GBASE-R) or 0:19 (for 100GBASE-R)

NOTE 2—Optional state (inside the dotted box) is only required to support EEE capability.

Figure 2-33. *Alignment marker lock state diagram*

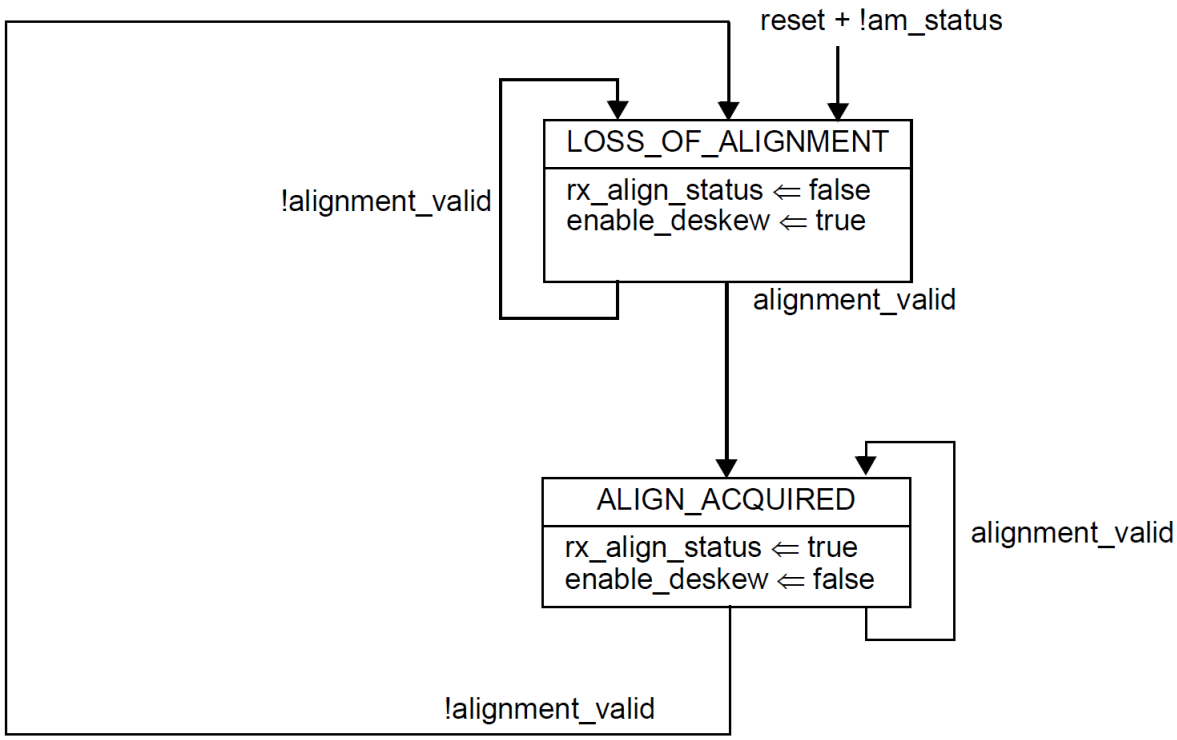


Figure 2-34. *PCS deskew state diagram.*

The PCS lanes can be received on different lanes of the service interface from which they were originally transmitted due to the Skew between the lanes and the multiplexing by the PMA. The RS-FEC transmit function shall order the PCS lanes according to the number of the PCS lane.

After all PCS lanes are aligned and deskewed, they are multiplexed together in the appropriate order to reconstruct the original stream of blocks and the alignment markers are removed from the data stream. Note that an alignment marker is always removed when the *am_lock* parameter is true for a given PCS lane even if it does not match the expected alignment marker value. The result of the repeated alignment marker errors in the *am_lock* parameter is set to false for a given PCS lane, but this happens until this is sufficient to remove the block in the alignment marker position.

For the optional EEE deep sleep capability, transitions between normal alignment markers and Rapid Alignment Markers (RAMs) result in changes in relative position and frequency of alignment markers. These transitions are detected by the Transmit LPI state diagram (see **Figure 2-35**), and this information is used by the alignment marker removal function to determine which 66-bit blocks are to be removed.

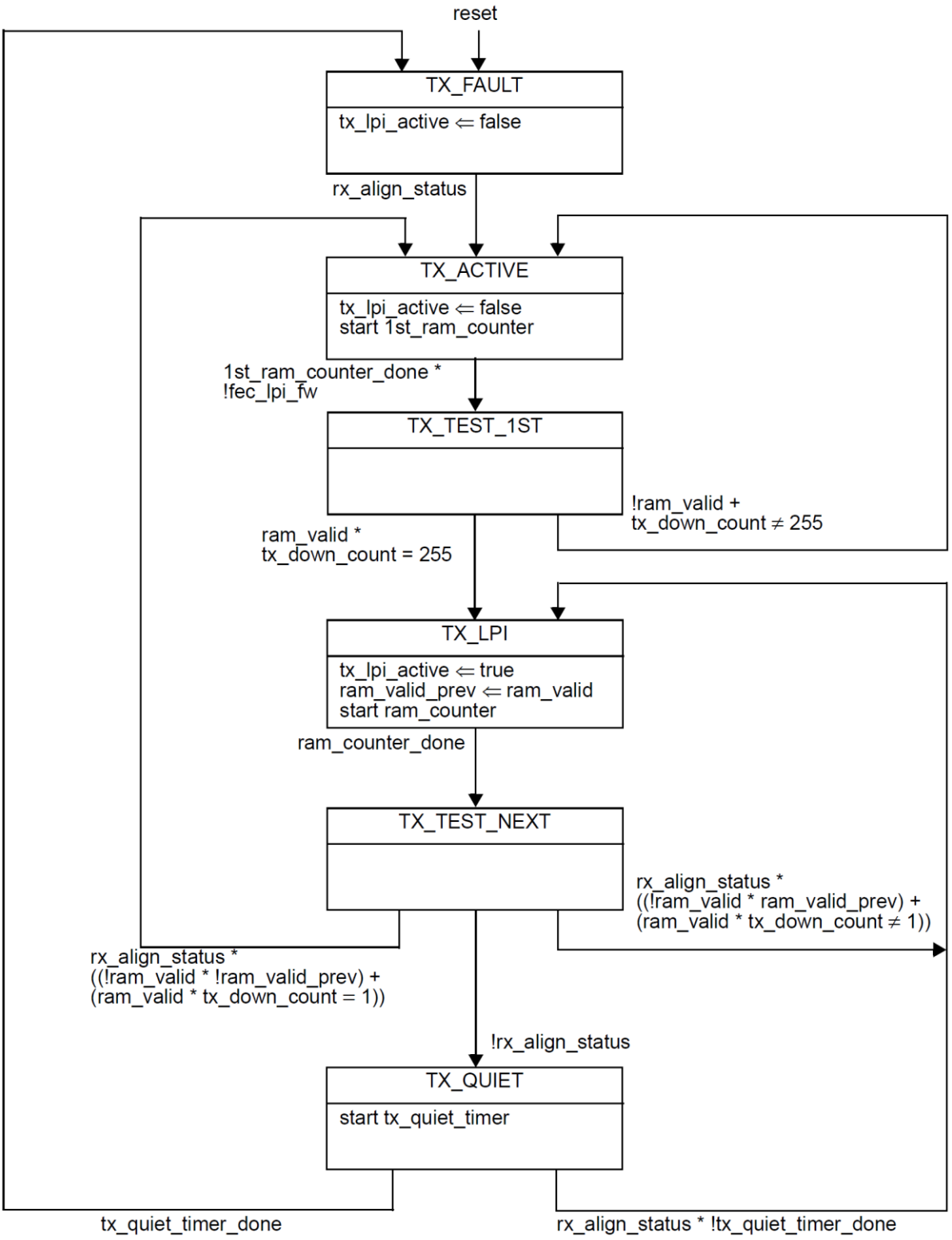


Figure 2-35. *Transmit LPI state diagram.*

As part of the alignment marker removal process, the BIP_3 field is compared to the calculated Bit Interleaved Parity (BIP) value for each PCS lane. If a Section **2.5** MDIO is implemented, then the appropriate BIP error counter register (registers 1.230 to

1.249) is incremented by one each time the calculated BIP value does not equal to the value received in the BIP_3 field. The bit error ratio in the data received from the local PCS is estimated by dividing the BIP block error ratio by a factor of 1 081 344.

NOTE: The data received from the local PCS is processed by the RS-FEC transmit function without any error correction.

The transcoder constructs a 257-bit block, *tx_scrambled<256:0>*, from a group of four 66-bit blocks, *tx_coded_j<65:0>* where *j*=0 to 3. For each group of the four 66-bit blocks, the *j*=3 corresponds to the most recently received block. The Bit 0 of each 66-bit block is the first bit received and it corresponds to the first bit of the synchronization header.

If for all j=0 to 3, tx_coded_j<0>=0 and tx_coded_j<1>=1, tx_xcoded<256:0> shall be constructed as follows:

a. *tx_xcoded<0>* = 1
b. *tx_xcoded<(64j+64):(64j+1)* > = *tx_coded_j<65:2>* for *j*=0 to 3

If for all *j*=0 to 3, *tx_coded_j<0>* ≠ *tx_coded_j<1>* (valid synchronization header) and for any *j*=0 to 3, *tx_coded_j<0>* = 1 and *tx_coded_j<1>* = 0, *tx_xcoded<256:0>* shall be constructed as follows:

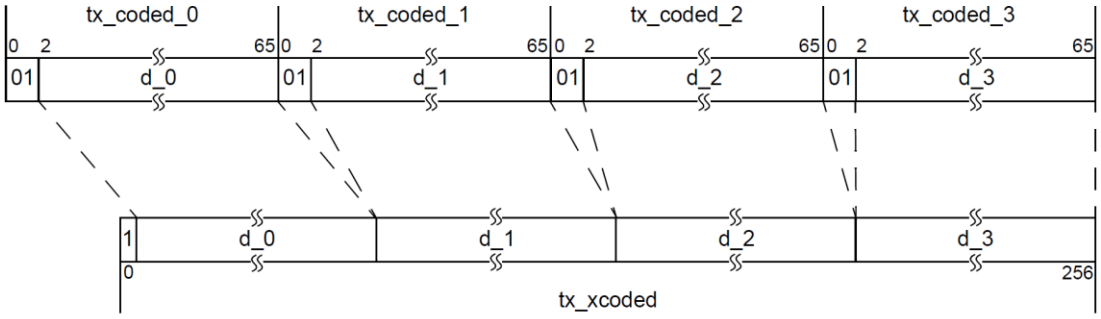
a. *tx_xcoded<0>* = 0
b. *tx_xcoded<j+1>* = *tx_coded_j<1>* for *j*=0 to 3
c. Let *c* be the smallest value of *j* such that *tx_coded_c<0>* = 1. Alternatively, the *tx_coded_c* is the first 66-bit control block that was received in the current group of four blocks.
d. Let *tx_payloads<(64j+63):64j>* = *tx_coded_j<65:2>* for *j*=0 to 3
e. Omit *tx_coded_c<9:6>*, which is the second nibble (based on transmission order) of the block type field for *tx_coded_c*, from *tx_xcoded* per the following expressions.
f. *tx_xcoded<(64c+8):5>* = *tx_payloads<(64c+3):0>*
g. *tx_xcoded<256:(64c+9)>* = *tx_payloads<255:(64c+8)>*

If for any *j*=0 to 3, *tx_coded_j<0>* = *tx_coded_j<1>* (invalid synchronization header), *tx_xcoded<256:0>* shall be constructed as follows:

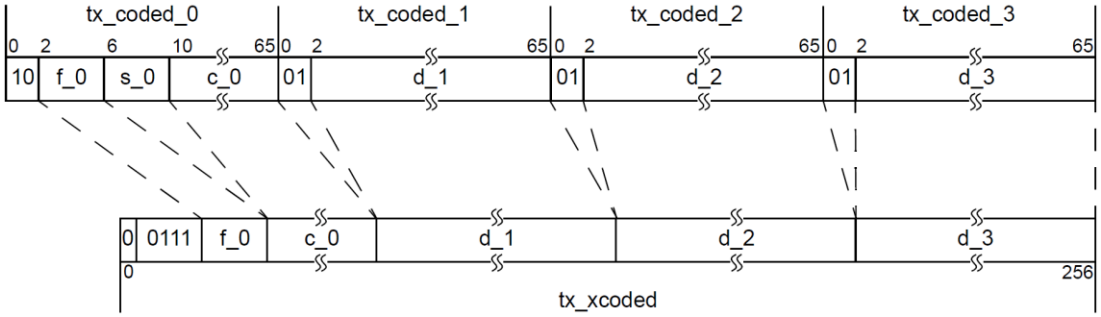
a. *tx_xcoded<0>* = 0
b. *tx_xcoded<j+1>* = 1 for *j*=0 to 3
c. Let *tx_payloads<(64j+63):64j>* = *tx_coded_j<65:2>* for *j*=0 to 3

d. Omit the second nibble (based on transmission order) of *tx_coded_0* per the following expressions.
    i. *tx_xcoded<8:5> = tx_payloads<3:0>*
    ii. *tx_xcoded<256:9> = tx_payloads<255:8>*

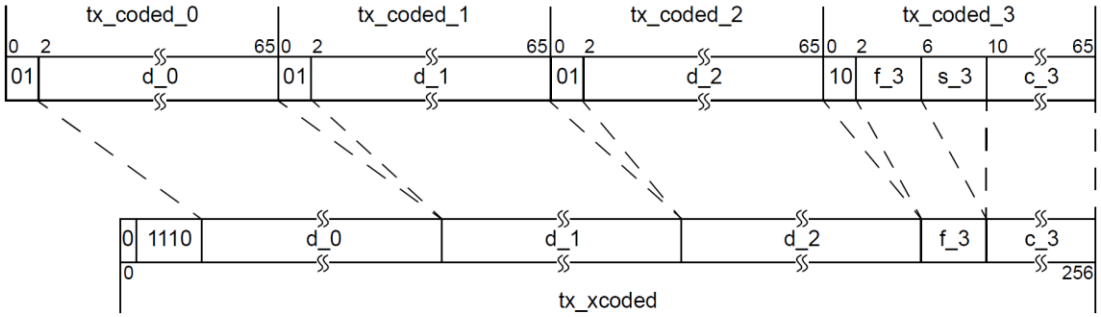
Several examples of the construction of *tx_xcoded<256:0>* are shown in **Figure 2-36**. In **Figure 2-36**, *d_j* indicates the *jth* 66-bit block contains only data octets, *c_j* indicates the *jth* 66-bit block contains one or more control characters, *f_j* indicates the first part of the block type field for 66-bit block *j*, and *s_j* indicates the second part of the block type field for 66-bit block *j*.



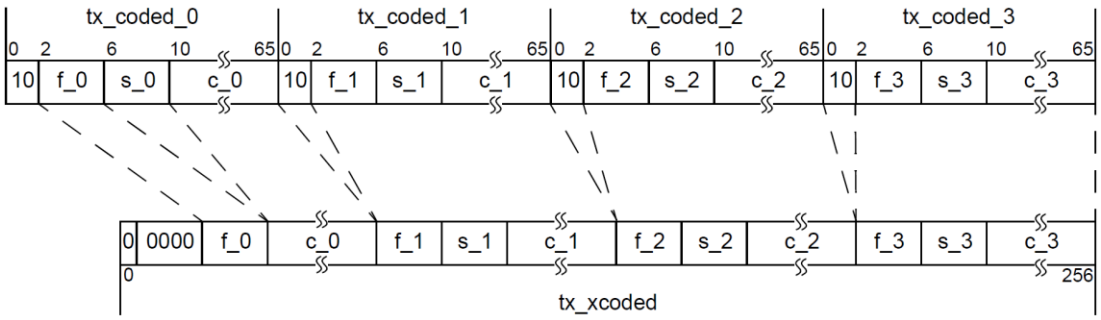
Example 1: All data blocks



Example 2: Control block followed by three data blocks



Example 3: Three data blocks followed by a control block



Example 4: All control blocks

Figure 2-36. *Examples of the construction of tx_xcoded.*

Finally, scramble the first 5 bits, based on transmission order, of *tx_xcoded<256:0>* to yield *tx_scrambled<256:0>* as follows:

a. Set *tx_scrambled<4:0>* to the result of the bit-wise exclusive-OR of the *tx_xcoded<4:0>* and *tx_xcoded<12:8>*.
b. Set *tx_scrambled<256:5>* to *tx_xcoded<256:5>*

For each 257-bit block, bit 0 is the first bit transmitted.

2.8.4.1.1 Reed-Solomon encoder

The RS-FEC sublayer uses a Reed-Solomon code, which operates over the Galois Field GF(2^{10}), where the symbol size is 10 bits. The encoder processes *k* message symbols to generate *2t* parity symbols, which are then appended to the message to produce a codeword of *n=k+2t* symbols. For the purposes of the IEEE standards, a particular Reed-Solomon code is denoted *RS(n,k)*.

When used to define a 100GBASE-SR4 PHY, the RS-FEC sublayer shall use the *RS(528,514*). Each k-symbol message corresponds to 20 blocks of 257-bit produced by the transcoder. Each code is based on the generating polynomial given by **Equation 2-5**.

$$g(x) = \prod_{j=0}^{2t-1}(c - \alpha^j) = g_{2t}x^{2t} + g_{2t-1}x^{2t-1} + \cdots + g_1x + g_0$$

Equation 2-5

In **Equation 2-5**, *α* is a primitive element of the finite field defined by the polynomial $x^{10}+x^3+1$. **Equation 2-6** defines the message polynomial m(x) whose coefficients are the message symbols m_{k-1} to m_0.

$$m(x) = m_{k-1}x^{n-1} + m_{k-2}x^{n-2} + \cdots + m_1x^{2t+1} + m_0x^{2t}$$

Equation 2-6

Each message symbol m_i is the bit vector $(m_{i,9}, m_{i,8}, \ldots, m_{i,1}, m_{i,0})$, which is identified with the element $m_{i,9}\alpha^9 + m_{i,8}\alpha^8 + \ldots + m_{i,1}\alpha + m_{i,0}$ of the finite field. The message symbols are consisted of the bits of the transcoded blocks *tx_scrambled* (including the mapped group of alignment markers when appropriate) such that bit 0 of the first transcoded block in the message (or *am_txmapped<0>*) is bit 0 of m_{k-1} and bit 256 of the last transcoded block in the message is bit 9 of m_0. The first symbol input at the encoder is the m_{k-1}.

Equation 2-7 defines the parity polynomial *p(x)*, whose coefficients are the parity symbols p_{2t-1} to p_0.

$$p(x) = p_{2t-1}x^{2t-1} + p_{2t-2}x^{2t-2} + \cdots + p_1x + p_0$$

Equation 2-7

The parity polynomial is obtained by the remainder from the division of m(x) by g(x). This is computed using the shift register implementation illustrated in **Figure 2-37**. The outputs of the delay elements are set to zero prior to the computation of the parity for a given message. After the symbol of the last message, m_0, is processed by the encoder, the outputs of the delay elements are the parity symbols for that message.

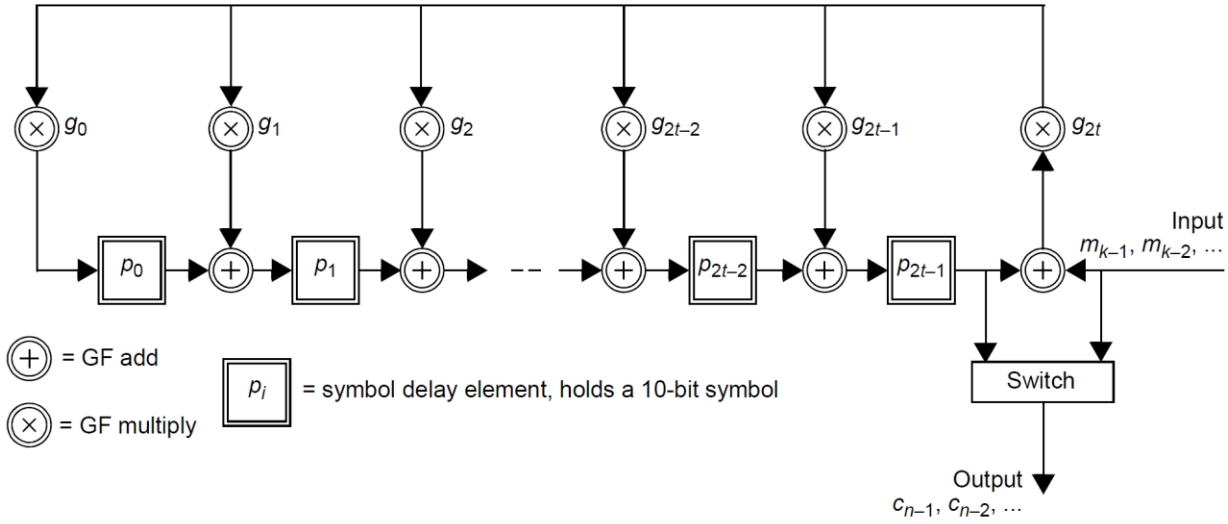


Figure 2-37. *Reed-Solomon encoder functional model*

Then, the codeword polynomial c(x) can be the sum of *m(x)* and *p(x)*, where the coefficient of the highest power of *x*, $c_{n-1} = m_{k-1}$ is transmitted first and the coefficient of the lowest power of *x*, $c_0 = p_0$ is transmitted last. The first bit transmitted from each symbol is the bit 0.

The coefficients of the generator polynomial for each code are shown in **Table 2-35**.

i	RS(528,514)	RS(544,514)	i	RS(528,514)	RS(544,514)	i	RS(528,514)	RS(544,514)
0	432	523	11	701	883	22		565
1	290	834	12	6	503	23		108
2	945	128	13	904	942	24		1
3	265	158	14	1	385	25		552
4	592	185	15		495	26		230
5	391	127	16		720	27		187
6	614	392	17		94	28		552
7	900	193	18		132	29		575
8	925	610	19		593	30		1
9	656	788	20		249			
10	32	361	21		282			

Table 2-35. *Coefficients of the generator polynomial g_i (decimal).*

Once the data has been encoded by Reed-Solomon, they shall be distributed to 4 FEC lanes, one 10-bit symbol at a time in a round robin distribution from the lowest to the highest numbered FEC lane. The distribution process is shown in **Figure 2-38**.

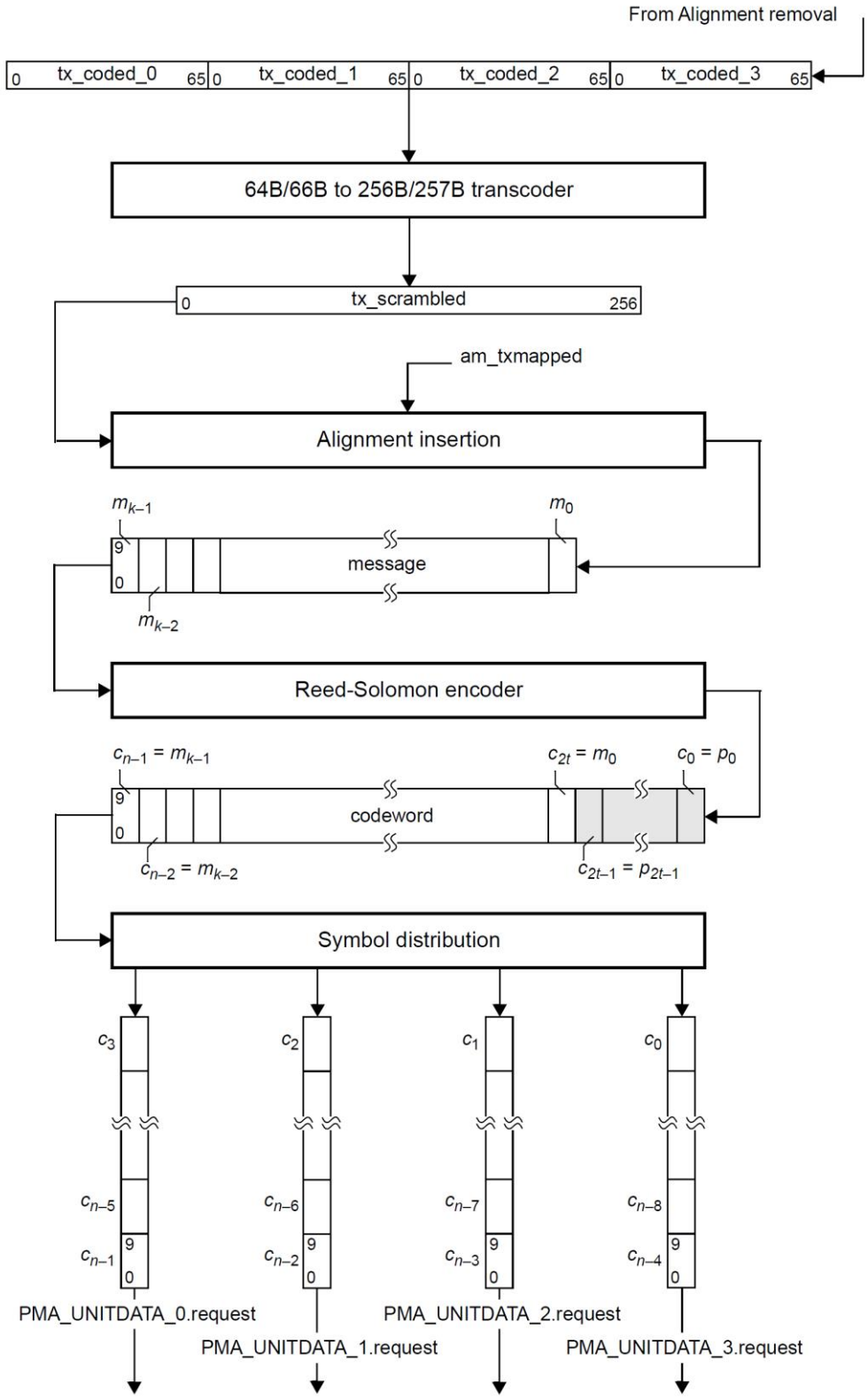


Figure 2-38. *Transmit bit ordering.*

The transmit bit ordering is illustrated in **Figure 2-38**.

2.8.4.2 Receive function

The RS-FEC receive function consists of 4 bit streams by concatenating the bits from each of the 4 *PMA:IS_UNITDATA_i.indication* primitives in the order they are received. It obtains lock to the alignment markers as specified by the FEC synchronization state diagram shown in **Figure 2-39**.

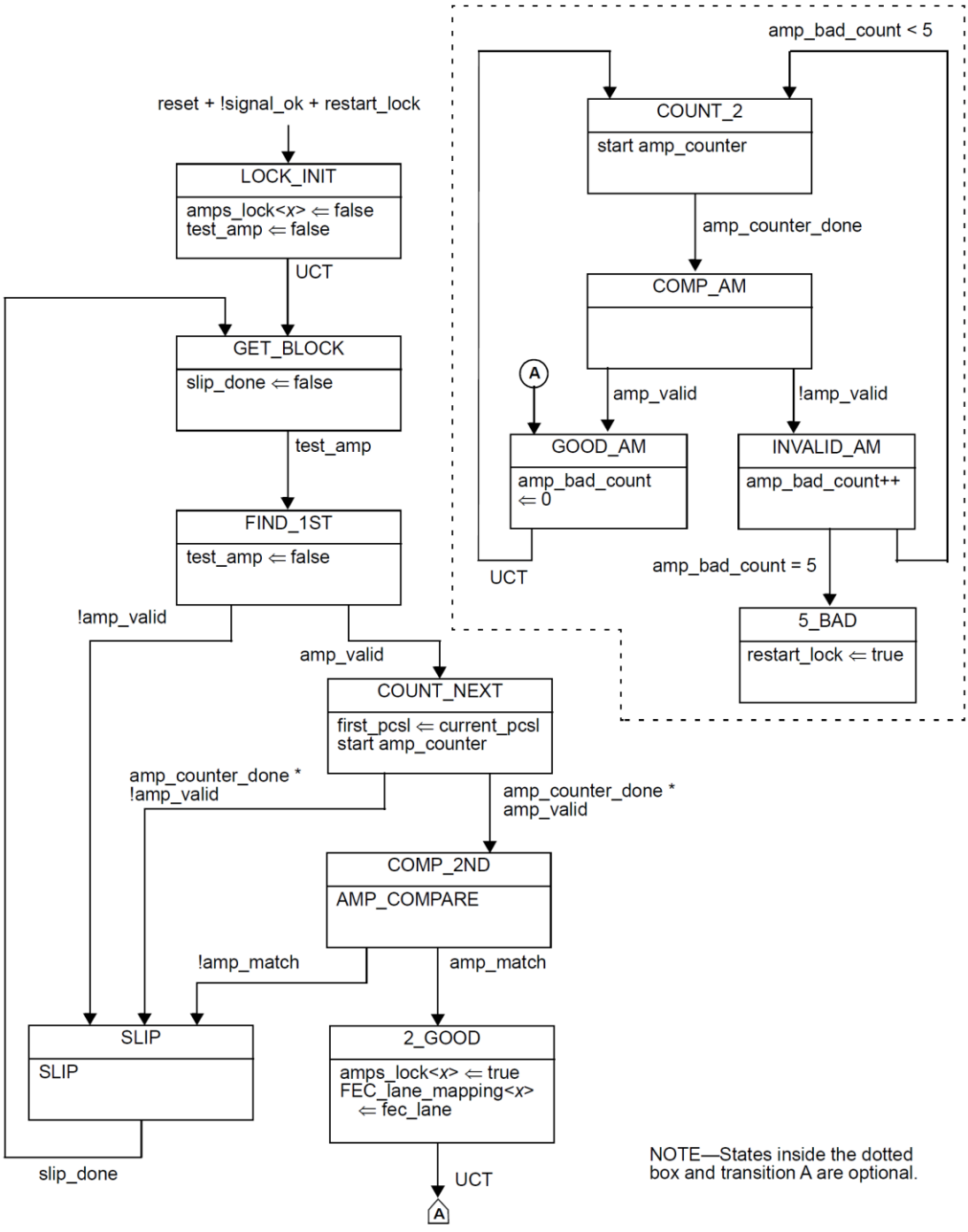


Figure 2-39. *FEC synchronization state diagram.*

After the alignment marker lock is achieved on all 4 lanes, all inter-lane Skew is removed as specified by the diagram of the FEC alignment state shown in **Figure 2-40**. The FEC receive function shall support a maximum Skew of 180ns between FEC lanes and a maximum Skew Variation of 4ns.

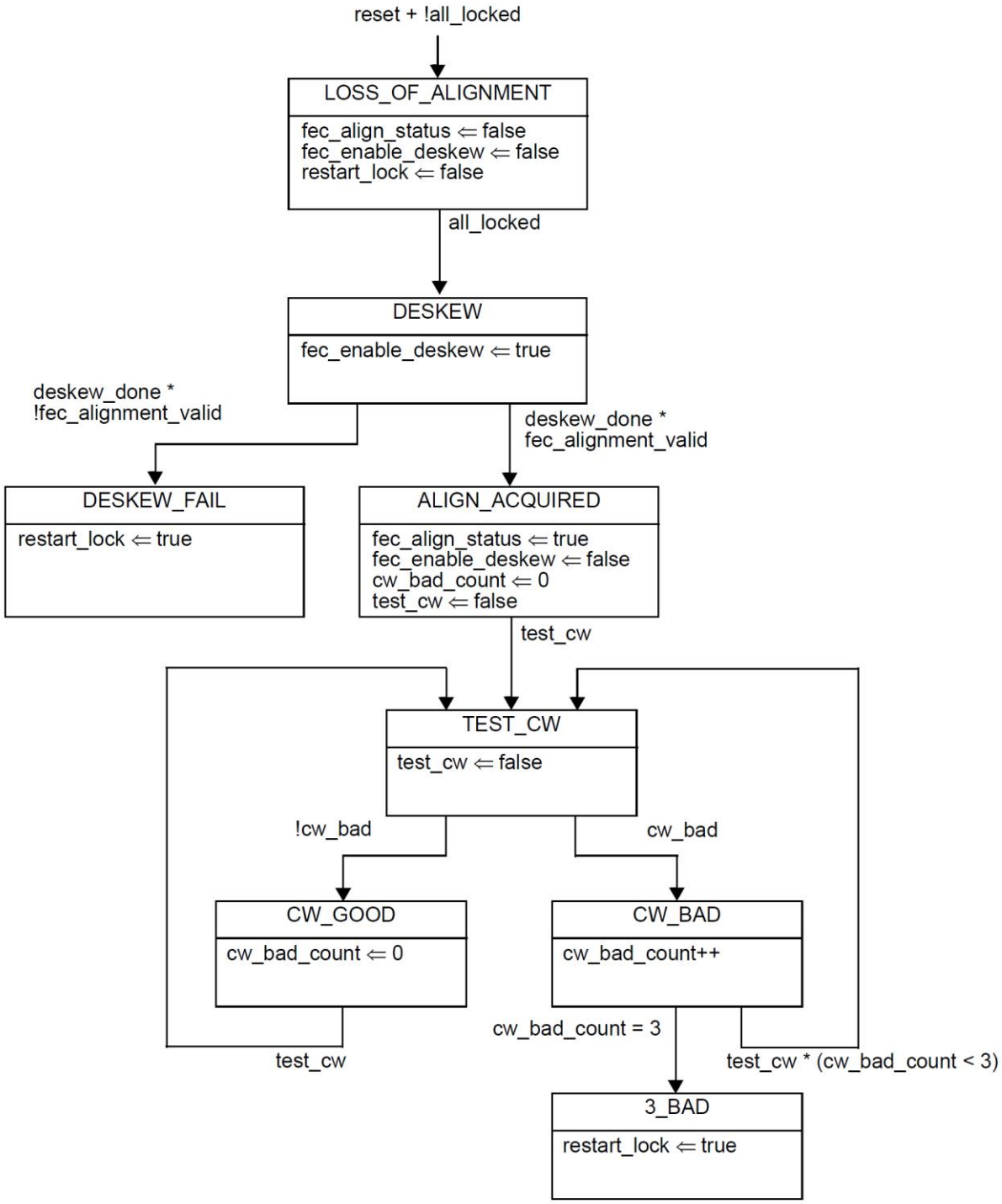


Figure 2-40. *FEC alignment state diagram.*

FEC lanes are received on different lanes of the service interface from which they were originally transmitted. The FEC receive function orders the FEC lanes according to the FEC lane number. The FEC lane number is set by the sequence of alignment markers that are mapped to each FEC lane.

After all FEC lanes are aligned, deskewed, and reordered, they are multiplexed together in the proper order to reconstruct the original stream of FEC codewords.

2.8.4.2.1 Reed-Solomon decoder

The Reed-Solomon decoder shall extract the message symbols from the codeword, shall correct them, and shall discard the parity symbols. The message symbols consist of 20 transcoded blocks *rx_scrambled*.

When used to form a 100GBASE-SR4 PHY, the RS-FEC sublayer corrects any combination of up to *t=7* symbol errors in a codeword. The RS-FEC sublayer shall also define when an errored codeword was not corrected. The probability that the decoder fails to indicate a codeword with *t+1* errors as uncorrected is not expected to exceed 10^{-6}. This limit is also expected to apply for *t+2* errors, *t+3* errors, and so on.

The Reed-Solomon decoder provides the error detection option without any option of error correction to reduce the delay contributed by the RS-FEC sublayer. This option is using the *FEC_bypass_correction_ability* variable, is enabled using the *FEC_bypass_correction_enable* variable and cannot be used when the RS-FEC sublayer is part of a 100GBASE-SR4 PHY.

The Reed-Solomon decoder defines the errors to the PCS sublayer by intentionally corrupting the synchronization headers of 66-bit block. When the decoder detects that a codeword contains errors or contains errors that were not corrected, it ensures that, for every other 257-bit block within the codeword starting with the first, the synchronization header for the first 66-bit block at the output of the 256B/257B to 64B/66B transcoder, *rx_coded_0<1:0>*, is set to 11. In addition, it shall ensure *rx_coded_0<1:0>* corresponding to the 6th 257-bit block and *rx_coded_3<1:0>* corresponding to the last (20th) 257-bit block in the codeword are set to 11. This causes the PCS to discard all frames 64 bytes and larger that are fully or partially within the codeword.

The Reed-Solomon decoder can optionally bypass the feature of error indication to reduce the delay provided by the RS-FEC sublayer. This option is indicated using the

FEC_bypass_indication_ability variable. When this option is provided, it is enabled using the *FEC_bypass_indication_enable* variable.

When *FEC_bypass_indication_enable* is used, additional error monitoring is performed by the RS-FEC sublayer to reduce the possibility that errors in a packet are not detected. The Reed-Solomon decoder counts the detected number of symbol errors of all four FEC lanes in the consecutive non-overlapping blocks of 8192 codewords. When the number of symbol errors in a block of 8192 codewords exceeds *K*, the Reed-Solomon decoder causes synchronization header *rx_coded<1:0>* of each subsequent 66-bit block that is delivered to the PCS to be assigned a value of 00 or 11 for a period of 60ms to 75ms. As a result, the PCS sets *hi_ber* = *true*, by limiting the processing of received packets. When Auto-Negotiation is supported and enabled, the *hi_ber* causes Auto-Negotiation to restart.

For the optional EEE deep sleep capability, the error monitor used when the *FEC_bypass_indication_enable* is disabled when *rx_lpi_active* = *true*. The next block of 8192 codewords defined by the error monitor starts on the codeword boundary, which follows the transition of *rx_lpi_active* from true to false.

When the RS-FEC sublayer is used for the 100GBASE-SR4 PHY, the threshold of the symbol error shall be K=417.

The first 1285 message bits in every 4096th codeword is the vector *am_rxmapped<1284:0>* where bit 0 is the first bit received. The specific codewords that use the 4096th vector are defined by the alignment lock and deskew function.

For the optional EEE deep sleep capability, transitions between normal alignment markers and Rapid Alignment Markers result in changes in the relative position and frequency of *am_rxmapped<1284:0>*. These transitions are defined by the diagram of the Receive LPI state (see **Figure 2-41**) and they are used by the alignment marker removal function to determine which bits are to be removed. When *rx_lpi_active* is true, the first 1285 message bits in every other codeword is the vector *am_rxmapped<1284:0>*.

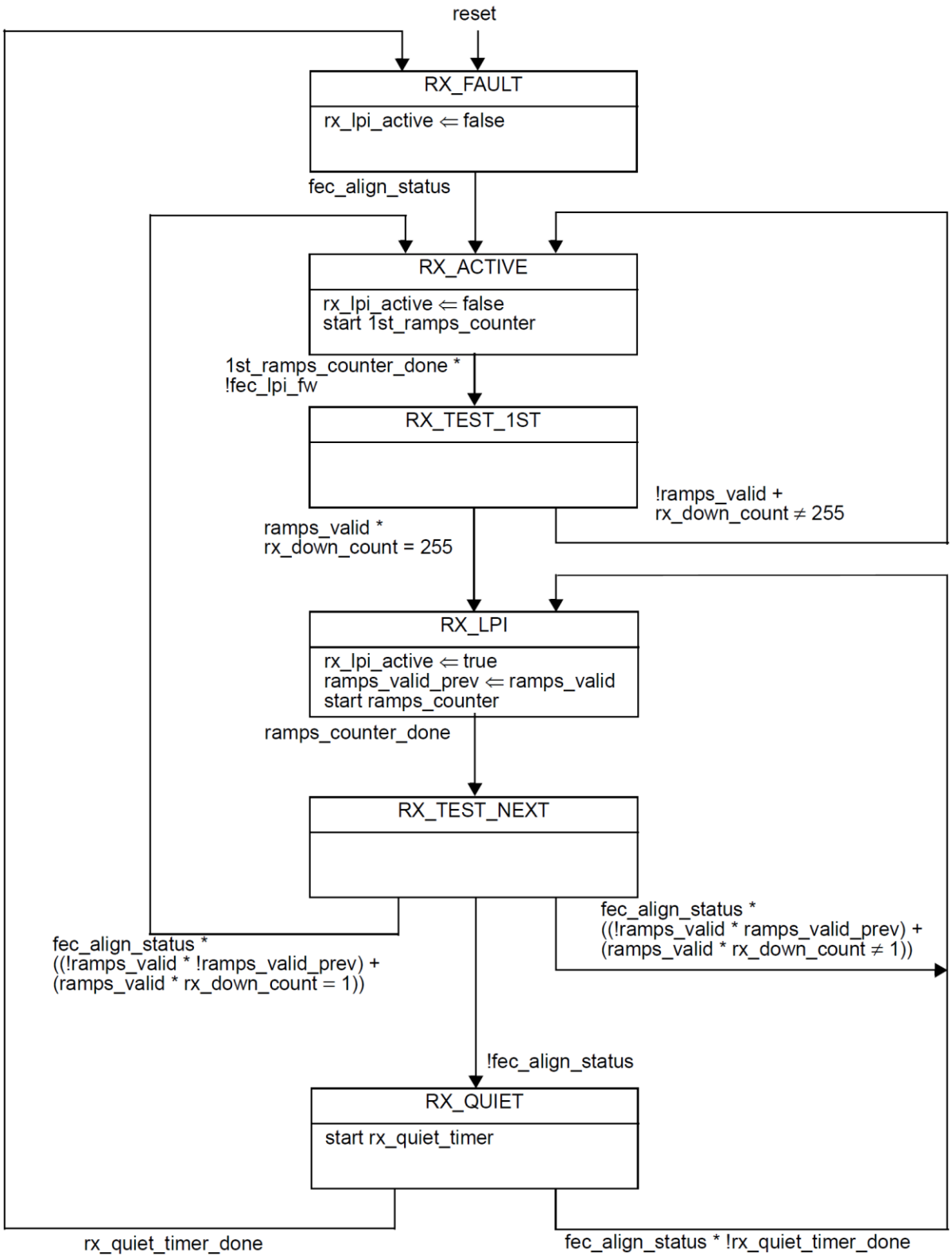


Figure 2-41. *Receive LPI state diagram*

The vector am_rxmapped shall be removed prior to transcoding.

2.8.4.2.2 256B/257B to 64B/66B transcoder

The transcoder extracts a group of four 66-bit blocks, *rx_coded_j<65:0>* where *j*=0 to 3, from each 257-bit block *rx_scrambled<256:0>*. Bit 0 of the 257-bit block is the first bit received.

First, descramble the first 5 bits, based on reception order, of *rx_scrambled<256:0>* to yield *rx_xcoded<256:0>* as follows.

a. Set *rx_xcoded<4:0>* to the result of the bit-wise exclusive-OR of the *rx_scrambled<4:0>* and *rx_scrambled<12:8>*.
b. Set *rx_xcoded<256:5>* to *rx_scrambled<256:5>*.

If *rx_xcoded<0>* is 1, *rx_coded_j<65:0>* for *j*=0 to 3 shall be derived as follows.

a. *rx_coded_j<65:2>* = *rx_xcoded<(64j+64):(64j+1)>* for *j*=0 to 3
b. *rx_coded_j<0>* = 0 and *rx_coded_j<1>* = 1 for all *j*=0 to 3

If *rx_xcoded<0>* is 0 and any *rx_xcoded<j+1* >= 0 for *j*=0 to 3, *rx_coded_j<65:0>* for *j*=0 to 3 shall be derived as follows.

a. Let *c* be the smallest value of *j* such that *rx_xcoded<j+1>* = 0. In other words, *rx_coded_c* is the first 66-bit control block in the resulting group of four blocks.
b. Let *rx_payloads* be a vector representing the payloads of the four 66-bit blocks. It is derived using the following expressions:
    i. *rx_payloads<(64c+3):0>* = *rx_xcoded<(64c+8):5>*
    ii. *rx_payloads<(64c+7):(64c+4)>* = 0000 (an arbitrary value that is later replaced by *s_c*)
    iii. *rx_payloads<255:(64c+8)>* = *rx_xcoded<256:(64c+9)>*
c. *rx_coded_j<65:2>* = *rx_payloads<(64j+63):64j>* for *j*=0 to 3
d. Let *f_c<3:0>* = *rx_coded_c<5:2>* be the scrambled first nibble (based on transmission order) of the block type field for *rx_coded_c*.
e. Descramble *f_c<3:0>* to yield *g<3:0>* per the following expression where **"^"** denotes the exclusive-OR operation. When *c*=0, *rx_coded_(c–1)* corresponds to *rx_coded_3* from the previous 257-bit block. *g<i>* = *f_c<i>* ^ *rx_coded_(c–1)<i+8>* ^ *rx_coded_(c–1)<i+27>* for *i*=0 to 3
f. The block type field may be uniquely identified by either its most or least significant nibble. Since *g<3:0>* is the least significant nibble of the block type field (per the transmission order), derive *h<3:0>* by cross-referencing to *g<3:0>*

using **Figure 2-42**. For example, if *g<3:0>* is *0xE* then *h<3:0>* is *0x1*. If no match to *g<3:0>* is found, *h<3:0>* is set to 0000.

g. If *rx_xcoded<j+1>*=0, *rx_coded_j<0>*=1 and *rx_coded_j<1>*=0 for *j*=0 to 3

h. If *rx_xcoded<j+1>*=1, *rx_coded_j<0>*=0 and *rx_coded_j<1>*=1 for *j*=0 to 3

i. If *h<3:0>* = 0000, *rx_coded_c<1>*=1 (invalidate synchronization header)

Input Data	Sync	Block Payload							
Bit Position: / Data Block Format:	0 1	2							65
$D_0\ D_1\ D_2\ D_3\ D_4\ D_5\ D_6\ D_7$	01	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7
Control Block Formats:		**Block Type Field**							
$C_0\ C_1\ C_2\ C_3\ C_4\ C_5\ C_6\ C_7$	10	0x1E	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7
$S_0\ D_1\ D_2\ D_3\ D_4\ D_5\ D_6\ D_7$	10	0x78	D_1	D_2	D_3	D_4	D_5	D_6	D_7
$O_0\ D_1\ D_2\ D_3\ Z_4\ Z_5\ Z_6\ Z_7$	10	0x4B	D_1	D_2	D_3	O_0	0x000_0000		
$T_0\ C_1\ C_2\ C_3\ C_4\ C_5\ C_6\ C_7$	10	0x87		C_1	C_2	C_3	C_4	C_5	C_6	C_7
$D_0\ T_1\ C_2\ C_3\ C_4\ C_5\ C_6\ C_7$	10	0x99	D_0		C_2	C_3	C_4	C_5	C_6	C_7
$D_0\ D_1\ T_2\ C_3\ C_4\ C_5\ C_6\ C_7$	10	0xAA	D_0	D_1		C_3	C_4	C_5	C_6	C_7
$D_0\ D_1\ D_2\ T_3\ C_4\ C_5\ C_6\ C_7$	10	0xB4	D_0	D_1	D_2		C_4	C_5	C_6	C_7
$D_0\ D_1\ D_2\ D_3\ T_4\ C_5\ C_6\ C_7$	10	0xCC	D_0	D_1	D_2	D_3		C_5	C_6	C_7
$D_0\ D_1\ D_2\ D_3\ D_4\ T_5\ C_6\ C_7$	10	0xD2	D_0	D_1	D_2	D_3	D_4		C_6	C_7
$D_0\ D_1\ D_2\ D_3\ D_4\ D_5\ T_6\ C_7$	10	0xE1	D_0	D_1	D_2	D_3	D_4	D_5		C_7
$D_0\ D_1\ D_2\ D_3\ D_4\ D_5\ D_6\ T_7$	10	0xFF	D_0	D_1	D_2	D_3	D_4	D_5	D_6	

Figure 2-42. *64B/66B block formats*

If rx_xcoded<0> is 0 and all rx_xcoded<j+1>=1 for j=0 to 3, rx_coded_j<65:0> for j=0 to 3 shall be derived as follows.

a. Set *c* = 0 and *h<3:0>* = 0000.

b. Let *rx_payloads* be a vector representing the payloads of the four 66-bit blocks. It is derived using the following expressions.

   i. *rx_payloads<(64c+3):0> = rx_xcoded<(64c+8):5>*

   ii. *rx_payloads<(64c+7):(64c+4)>* = 0000 (an arbitrary value that is later replaced by *s_c*)

   iii. *rx_payloads<255:(64c+8)> = rx_xcoded<256:(64c+9)>*

c. *rx_coded_j<65:2> = rx_payloads<(64j+63):(64j)>* for *j*=0 to 3

d. *rx_coded_j<0>*=0 and *rx_coded_j<1>*=0 for *j*=0 and 2

e. *rx_coded_j<0>*=1 and *rx_coded_j<1>*=1 for *j*=1 and 3

If *rx_xcoded<0>* is 0, scramble *h<3:0>* to yield *s_c<3:0>* and assign it to *rx_coded_c* per the following expressions.

a. *s_c<i> = h<i> ^ rx_coded_(c–1)<i+12> ^ rx_coded_(c–1)<i+31>* for *i*=0 to 3

b. *rx_coded_c<9:6> = s_c<3:0>*

The 66-bit blocks are transmitted in order from *j*=0 to 3. Bit 0 of each block is the first bit transmitted.

After the data has been transcoded, it shall be distributed to multiple PCS lanes, one 66-bit block at a time in a round robin distribution from the lowest to the highest numbered PCS lanes. The distribution process is shown in **Figure 2-43**.

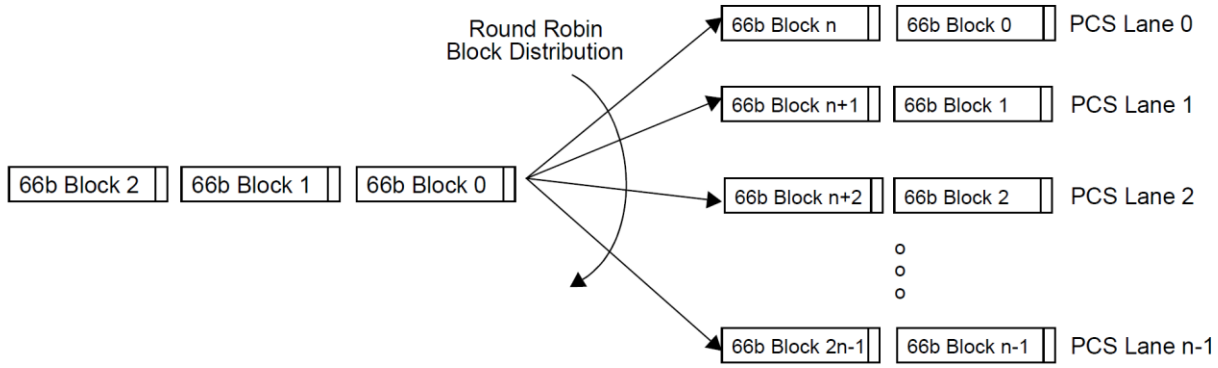


Figure 2-43. *PCS Block distribution.*

2.9 Physical Medium Dependent (PMD) sublayer and medium, type 100GBASE-SR4

This Section specifies the 100GBASE-SR4 PMD together with the multimode fiber medium. The PMD sublayer provides a point-to-point 100 Gb/s Ethernet link over four pairs of multimode fiber, up to at least 100m. When forming a complete PHY, a PMD shall be connected to the appropriate PMA as shown in **Table 2-36**, to the medium through the MDI and optionally with the management functions that may be accessible through the management interface, or equivalent.

Associated Section	100GBASE-SR4
2.3 - RS	Required
2.3 - CGMII [a]	Optional
2.4 - PCS for 100GBASE-R	Required
2.4 - PMA for 100GBASE-R	Required
2.8 - RS-FEC [b]	Required
Annex H - CAUI-10	Optional

Annex I - Chip-to-module CAUI-10 [c]	Optional
Annex K - CAUI-4	Optional
Annex L - Chip-to-module CAUI-4	Optional
Annex F - Energy Efficient Ethernet	Optional
[a] The CGMII is an optional interface. [b] The option to bypass the Section **2.8** RS-FEC correction function is not supported. [c] The RS-FEC sublayer of this option shall be within the module. See Section **2.8.2**.	

Table 2-36. *PHY sections associated with the 100GBASE-SR4 PMD.*

Figure 2-44 shows the relationship to the ISO/IEC OSI reference model of the PMD and MDI (shown shaded) with other sublayers. 100 Gb/s Ethernet is introduced in Section **2.2** and the purpose of each PHY sublayer is summarized in Section **2.3**. 100GBASE-SR4 PHYs with the optional EEE fast wake capability can enter the LPI mode to conserve energy during periods of low link utilization (see **Annex F**). The EEE deep sleep mode is not supported.

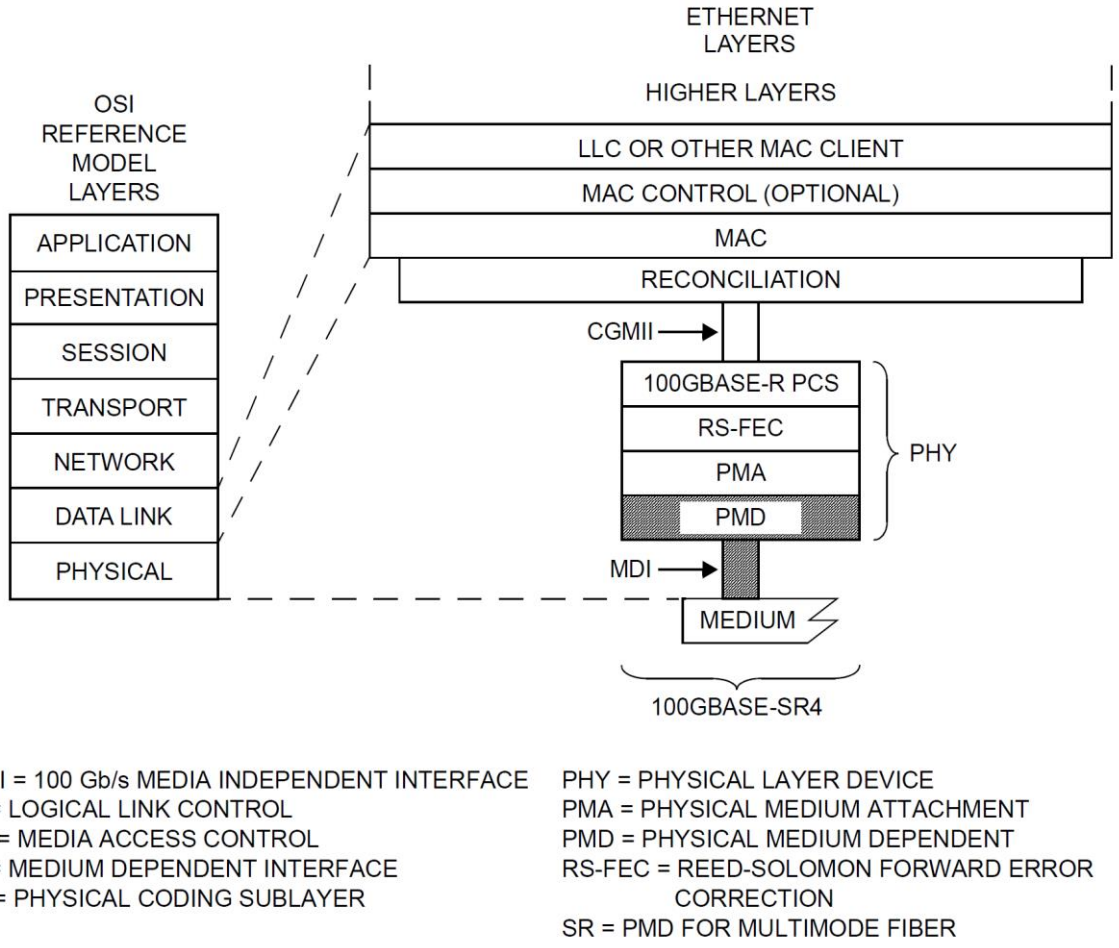


Figure 2-44. *100GBASE-SR4 PMD relationship of the ISO/IEC OSI reference model with the IEEE 802.3 Ethernet model.*

The BER shall be less than 5×10^{-5} provided that the error statistics are sufficiently random that this results in a frame loss ratio[1] of less than 6.2×10^{-10} for 64-octet frames with minimum interpacket gap according to Section **2.8**.

If the error statistics are not random enough to meet this requirement, then the BER shall be less than that required to give a frame loss ratio of less than 6.2×10^{-10} for 64-octet frames with minimum interpacket gap when processed according to Section **2.8**.

The PMD service interface supports the exchange of the encoded data between the PMA entity above the PMD, and the PMD entity. The PMD translates the encoded data to and from the signals of each specified medium.

The PMD service interface as an instance of the inter-sublayer service interface has the following primitives:

- *PMD:IS_UNITDATA_i.request*
- *PMD:IS_UNITDATA_i.indication*
- *PMD:IS_SIGNAL.indication*

The 100GBASE-SR4 PMD has four parallel bit streams, from *i* = 0 to 3.

In the transmit direction, the PMA continuously sends four parallel bit streams to the PMD, one per lane, each at a nominal signaling rate of 25.78125GBd. The PMD converts these bit streams into the appropriate signals on the MDI.

In the receive direction, the PMD continuously sends four parallel bit streams to the PMA, which correspond to the signals received from the MDI, one per lane, each at a nominal signaling rate of 25.78125GBd.

The *SIGNAL_DETECT* parameter defined in this Section maps the *SIGNAL_OK* parameter in the *PMD:IS_SIGNAL.indication(SIGNAL_OK)* inter-sublayer service primitive.

The *SIGNAL_DETECT* parameter takes one of the two values: *OK* or *FAIL*. When the *SIGNAL_DETECT* = *FAIL*, then the *rx_bit* parameters are undefined.

NOTE: The *SIGNAL_DETECT* = *OK* does not guarantee that the *rx_bit* parameters are good. It is possible for a poor-quality link to provide sufficient light for a *SIGNAL_DETECT* = *OK* indication and still not meet the BER defined in Section **2.9**.

[1] *frame loss ratio* is the number of transmitted frames not received as valid by the MAC divided by the total number of transmitted frames.

The sum of the transmit and receive delays at one end of the link comprised of 100GBASE-SR4 PMD including 2m of fiber in one direction shall be no more than 2048bit times (4 *pause_quanta* or 20.48ns). In Section **2.2.1** are defined the description of the overall system delay constraints, the definitions of the bit times and the *pause_quanta*.

The Skew (relative delay) between the lanes shall be within the limits, so that the information on the lanes can be reassembled by the RS-FEC sublayer. Skew and Skew Variation are defined in Section **2.2.2** and specified at the points SP0 to SP7 shown in **Figure 2-7**.

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

For more information on Skew and Skew Variation see Section **2.2.2**. The measurements of Skew and Skew Variation are defined in Section **2.6.4**.

2.9.1 PMD functional specifications

The 100GBASE-SR4 PMD performs the Transmit and Receive functions, which transmit data from the PMD service interface to the MDI.

The PMD block diagram is shown in **Figure 2-45**. The optical transmit signal is defined at the output end of a multimode fiber patch cord (TP2), between 2m and 5m in length. Unless otherwise specified, all transmitter measurements and tests defined in Section **2.9.4** are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see Section **2.9.7.1**). Unless otherwise specified, all receiver measurements and tests defined in Section **2.9.4** are made at TP3.

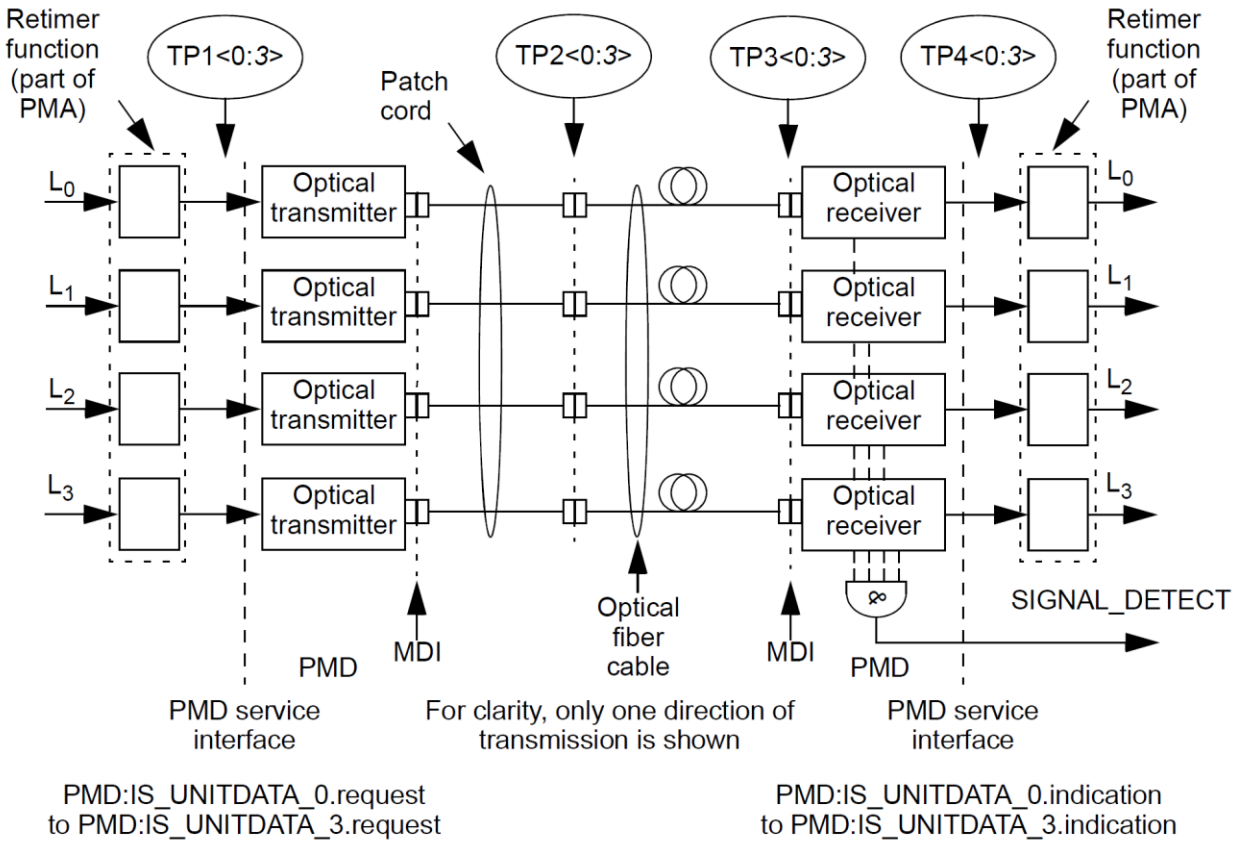


Figure 2-45. *Block diagram for 100GBASE-SR4 transmit/receive paths.*

TP1<0:3> and TP4<0:3> are informative reference points that may be useful to implementers for testing components (these test points will not typically be accessible in an implemented system).

2.9.1.1 PMD transmit function

The PMD Transmit function shall convert the four bit streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_3.request* into four separate optical signal streams. The four optical signal streams shall then be delivered to the MDI, which contains four parallel light paths for transmit, according to the transmit optical specifications of this Section. The *tx_bit* = one corresponds to the higher optical power level in each signal stream.

2.9.1.2 PMD receive function

The PMD Receive function shall convert the four parallel optical signal streams received from the MDI into separate bit streams for delivery to the PMD service interface using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_3.indication*, all according to the receive optical specifications of

this Section. The *rx_bit* = one correspond to the higher optical power level in each signal stream.

2.9.1.3 PMD global signal detect function

The function of PMD global signal detect reports the state of the *SIGNAL_DETECT* parameter through the PMD service interface. The *SIGNAL_DETECT* parameter is continuously signaled, while the message of *PMD:IS_SIGNAL.indication* is generated when a change occurs in the value of *SIGNAL_DETECT*. The *SIGNAL_DETECT* parameter maps the *SIGNAL_OK* parameter in the primitives of the inter-sublayer service interface.

SIGNAL_DETECT is a global indicator, which defines the presence of optical signals on all four lanes. The value of the *SIGNAL_DETECT* parameter shall be generated according to the conditions defined in **Table 2-37**. The PMD receiver is not necessary to verify if a compliant 100GBASE-SR4 signal is being received. This Section does not provide any response time requirements for the generation of the *SIGNAL_DETECT* parameter.

Receive conditions	SIGNAL_DETECT value
For any lane; Average optical power at TP3 ≤ -30 dBm	FAIL
For all lanes, [(Optical power at TP3 ≥ average receive power, each lane (min) in **Table 2-40**) AND (Compliant 100GBASE-SR4 signal input)]	OK
All other conditions	Unspecified

Table 2-37. *SIGNAL_DETECT value definition.*

The *SIGNAL_DETECT* parameter requires all implementations to provide adequate margin between the input optical power level, where the *SIGNAL_DETECT = OK,* and the inherent noise level of the PMD, including the effects of crosstalk, power supply noise, etc.

The various implementations of the *SIGNAL_DETECT* function are permitted, including the implementations that generate the values *of the SIGNAL_DETECT* parameter in response to the amplitude of the modulation of the optical signal, which have the average optical power of the modulated optical signal.

2.9.1.4 PMD lane-by-lane signal detect function

For the implementation of the MDIO, each *PMD_signal_detect_i*, where *i* represents the lane number in the range 0 to 3, shall be continuously set in response to the magnitude of the optical signal of its associated lane, according to the requirements of **Table 2-37**.

2.9.2 Lane assignments

There are no lane assignments (within each group of transmit or receive lanes) for 100GBASE-SR4. While the PMD is going to map the electrical lane i to the optical lane i and vice versa, the physical ordering of the lanes is not necessary to be defined, since the RS-FEC sublayer is receiving the lanes in any arrangement. The positioning of the transmit and receive lanes at the MDI is specified in Section **2.9.7.2**.

2.9.3 PMD to MDI optical specifications for 100GBASE-SR4

The operating range for the 100GBASE-SR4 PMD is defined in **Table 2-38**. A 100GBASE-SR4 compliant PMD operates on 50/125μm multimode fibers, type A1a.2 (OM3) or type A1a.3 (OM4), according to the specifications defined in **Table 2-45**. A PMD that exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., a 100GBASE-SR4 PMD operating at 120m meets the operating range requirement of 0.5m to 100m).

PMD type	Required operating range [a]
100GBASE-SR4	0.5m to 70m for OM3
	0.5m to 100m for OM4
[a] The RS-FEC correction function cannot be bypassed for any operating distance.	

Table 2-38. *100GBASE-SR4 operating range.*

2.9.3.1 100GBASE-SR4 transmitter optical specifications

Each lane of a 100GBASE-SR4 transmitter shall meet the specifications in **Table 2-39** per the definitions in Section **2.9.4**.

Description	Value	Unit
Signaling rate, each lane (range)	25.78125 ± 100 ppm	GBd
Center wavelength (range)	840 to 860	nm
RMS spectral width [a] (max)	0.6	nm
Average launch power, each lane (max)	2.4	dBm
Average launch power, each lane (min)	-8.4	dBm
Optical Modulation Amplitude (OMA), each lane (max)	3	dBm

Optical Modulation Amplitude (OMA), each lane (min) [b]	-6.4	dBm
Launch power in OMA minus TDEC (min)	-7.3	dBm
Transmitter and dispersion eye closure (TDEC), each lane (max)	4.3	dB
Average launch power of OFF transmitter, each lane (max)	-30	dBm
Extinction ratio (min)	2	dB
Optical return loss tolerance (max)	12	dB
Encircled flux [c]	≥ 86% at 19μm ≤ 30% at 4.5μm	
Transmitter eye mask definition, {X1, X2, X3, Y1, Y2, Y3} Hit ratio 1.5 × 10^{-3} hits per sample	{0.3, 0.38, 0.45, 0.35, 0.41, 0.5}	
[a] The RMS spectral width is the standard deviation of the spectrum. [b] Even if the TDEC < 0.9dB, the OMA (min) must exceed this value. [c] If the 50μm fiber type A1a.2 or type A1a.3 is measured in accordance with IEC 61280-1-4.		

Table 2-39. *100GBASE-SR4 transmit characteristics.*

2.9.3.2 100GBASE-SR4 receive optical specifications

Each lane of a 100GBASE-SR4 receiver shall meet the specifications in **Table 2-40** per the definitions in Section **2.9.4**.

Description	Value	Unit
Signaling rate, each lane (range)	25.78125 ± 100 ppm	GBd
Center wavelength (range)	840 to 860	nm
Damage threshold [a] (min)	3.4	dBm
Average receive power, each lane (max)	2.4	dBm
Average receive power, each lane [b] (min)	-10.3	dBm
Receive power, each lane (OMA) (max)	3	dBm
Receiver reflectance (max)	-12	dB
Stressed receiver sensitivity (OMA), each lane [c] (max)	-5.2	dBm
Conditions of stressed receiver sensitivity test: [d]		
Stressed eye closure (SEC), lane under test	4.3	dB
Stressed eye J2 Jitter, lane under test	0.39	UI
Stressed eye J4 Jitter, lane under test (max)	0.53	UI
OMA of each aggressor lane	3	dBm
Stressed receiver eye mask definition, {X1, X2, X3, Y1, Y2, Y3} Hit ratio 5 × 10^{-5} hits per sample	{0.28, 0.5, 0.5, 0.33, 0.33, 0.4}	

[a] The receiver shall tolerate without damage, continuous exposure to an optical input signal having the average power level on one lane. The receiver is not necessary to operate correctly at this input power.

[b] Average receive power, each lane (min) is informative and not the principal indicator of signal strength.

[c] Measured with conformance test signal at TP3 (see Section **2.9.4.9**) for the BER specified in Section **2.9**.

[d] These test conditions are for the measurement of the stressed receiver sensitivity. They are not characteristics of the receiver.

Table 2-40. *100GBASE-SR4 receive characteristics.*

2.9.3.3 100GBASE-SR4 illustrative link power budget

An illustrative power budget and penalties for 100GBASE-SR4 channels are shown in **Table 2-41**.

Parameter	OM3	OM4	Unit
Effective modal bandwidth at 850 nm [a]	2000	4700	MHz·km
Power budget (for max TDEC)	8.2		dB
Operating distance	0.5 to 70	0.5 to 100	m
Channel insertion loss [b]	1.8	1.9	dB
Allocation for penalties [c] (for max TDEC)	6.3		dB
Additional insertion loss allowed	0.1	0	dB

[a] Per IEC 60793-2-10.

[b] The channel insertion loss is calculated using the maximum distance specified in **Table 2-38** and cabled optical fiber attenuation of 3.5dB/km at 850nm plus an allocation for connection and splice loss given in Section **2.9.7**.

[c] Link penalties are used for the link budget calculations.

Table 2-41. *100GBASE-SR4 illustrative link power budget.*

2.9.4 Definition of optical parameters and measurement methods

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified.

While compliance must be achieved in normal operation, specific test patterns are defined for measurement consistency and the measurement of some parameters. **Table 2-42** gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subsections in which each parameter is defined. Any of the test patterns given for a particular test in **Table 2-42** may be used

to perform that test. The Pattern 3 is more demanding than Pattern 5 (which is the same or more demanding than other 100GBASE-R bit streams), so an item that is compliant using Pattern 5 is considered compliant even if it does not comply with the required limit using Pattern 3. The test patterns used in this Section are shown in **Table 2-43**.

Parameter	Pattern	Related Section
Wavelength, spectral width	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.1**
Average optical power	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.2**
Optical modulation amplitude (OMA)	Square wave or 4	**2.9.4.3**
Transmitter and dispersion eye closure (TDEC)	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.4**
Extinction ratio	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.7**
Transmitter optical waveform	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.8**
Stressed receiver sensitivity	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.9**
Stressed eye closure (SEC), calibration	3, 5 or valid 100GBASE-SR4 signal	**2.9.4.9**

Table 2-42. *Test-pattern definitions and related sections.*

Pattern	Pattern description
Square wave	Square wave (8 ones, 8 zeros)
3	PRBS31
4	PRBS9
5 [a]	RS-FEC encoded scrambled idle
[a] The pattern encoded by Section **2.8** RS-FEC for 100GBASE-SR4.	

Table 2-43. *Test patterns.*

The stressed receiver sensitivity is defined for an interface at the BER specified in Section **2.9**. The interface BER is the average of the four BERs of the receive lanes, when they are stressed.

Measurements with Pattern 3 (PRBS31) allow lane-by-lane BER measurements. Measurements with Pattern 5 (RS-FEC encoded scrambled idle) give the interface BER if all lanes are stressed at the same time. If each lane is stressed in turn, the BER is diluted by the three unstressed lanes, and the BER for that stressed lane alone must be calculated, e.g., by multiplying by four if the unstressed lanes have low BER. In stressed receiver sensitivity measurements, unstressed lanes can be created by setting the power at the receiver under test adequately above its sensitivity and/or not stressing those lanes with ISI and jitter, or by other means. Each receive lane is stressed in turn, while all others are operated. All aggressor lanes are operated as specified. The interface BER is calculated as the average of the BERs of all the lanes when stressed.

Where not otherwise specified, the maximum amplitude of the OMA or VMA is used for a particular situation, and also the minimum transition time is used for the counter-propagating lanes. Other test methods that generate equivalent results can be used. While the lanes in a particular direction can share a common clock, the Tx and Rx directions are not synchronous to each other. If Pattern 3 is used for the lanes not under test using a common clock, there is at least a delay of 31 UI between the PRBS31 patterns on one lane and any other lane.

2.9.4.1 Center wavelength and spectral width

The center wavelength and RMS spectral width of each optical lane shall be within the range given in **Table 2-40** if measured per IEC 61280-1-3. The lane under test is modulated using one of the test patterns specified in **Table 2-42**.

2.9.4.2 Average optical power

The average optical power of each lane shall be within the limits given in **Table 2-40** if measured using the methods defined in IEC 61280-1-1. The average optical power is measured using the test pattern defined in **Table 2-42**.

2.9.4.3 Optical Modulation Amplitude (OMA)

OMA shall be within the limits given in **Table 2-40** for measurement with a square wave (8 ones, 8 zeros) test pattern for measurement with a PRBS9 test pattern, with

the exception that each optical lane is tested individually. See Section **2.9.4** for test pattern information.

2.9.4.4 Transmitter and dispersion eye closure (TDEC)

TDEC of each lane shall be within the limits given in **Table 2-40** if measured using the methods specified in Sections **2.9.4.5** and **2.9.4.6**.

TDEC is a measure of each optical transmitter's vertical eye closure, which is based upon vertical histogram data from an eye diagram measured through an optical to electrical converter (O/E) with a bandwidth equivalent to a combined reference receiver and a worst-case optical channel. **Table 2-42** specifies the test patterns to be used for measurement of TDEC.

2.9.4.5 TDEC conformance test setup

A block diagram for the TDEC conformance test is shown in **Figure 2-46**. Other measurement implementations can be used with suitable calibration.

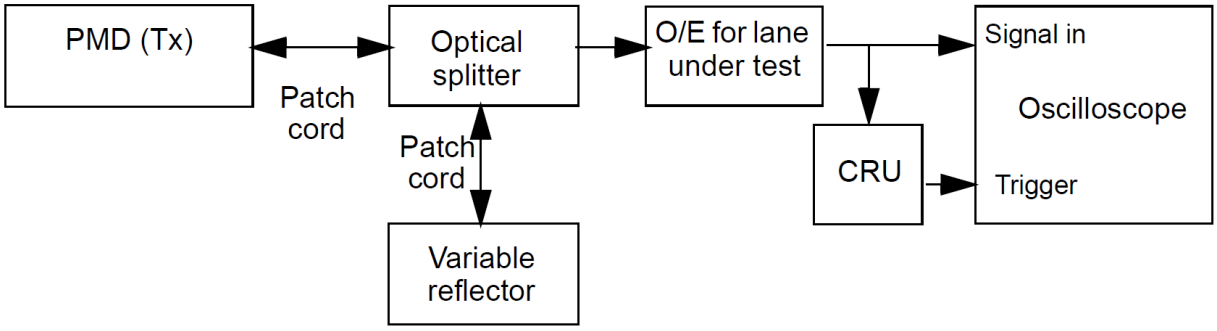


Figure 2-46. *TDEC conformance test block diagram.*

Each optical lane shall be tested individually with all other lanes in operation. The optical splitter and variable reflector are adjusted so that each transmitter is tested with an optical return loss of 12dB.

The combination of the O/E and the oscilloscope used to measure the optical waveform has a fourth-order Bessel-Thomson filter response with a bandwidth of 12.6GHz. Compensation can be made for all deviations from the ideal fourth-order Bessel-Thomson response.

The clock recovery unit (CRU) has a corner frequency of 10MHz and a slope of 20dB/decade.

2.9.4.6 TDEC measurement method

The oscilloscope is set up to accumulate samples of the optical eye diagram for the transmitter under test, as illustrated in **Figure 2-47**.

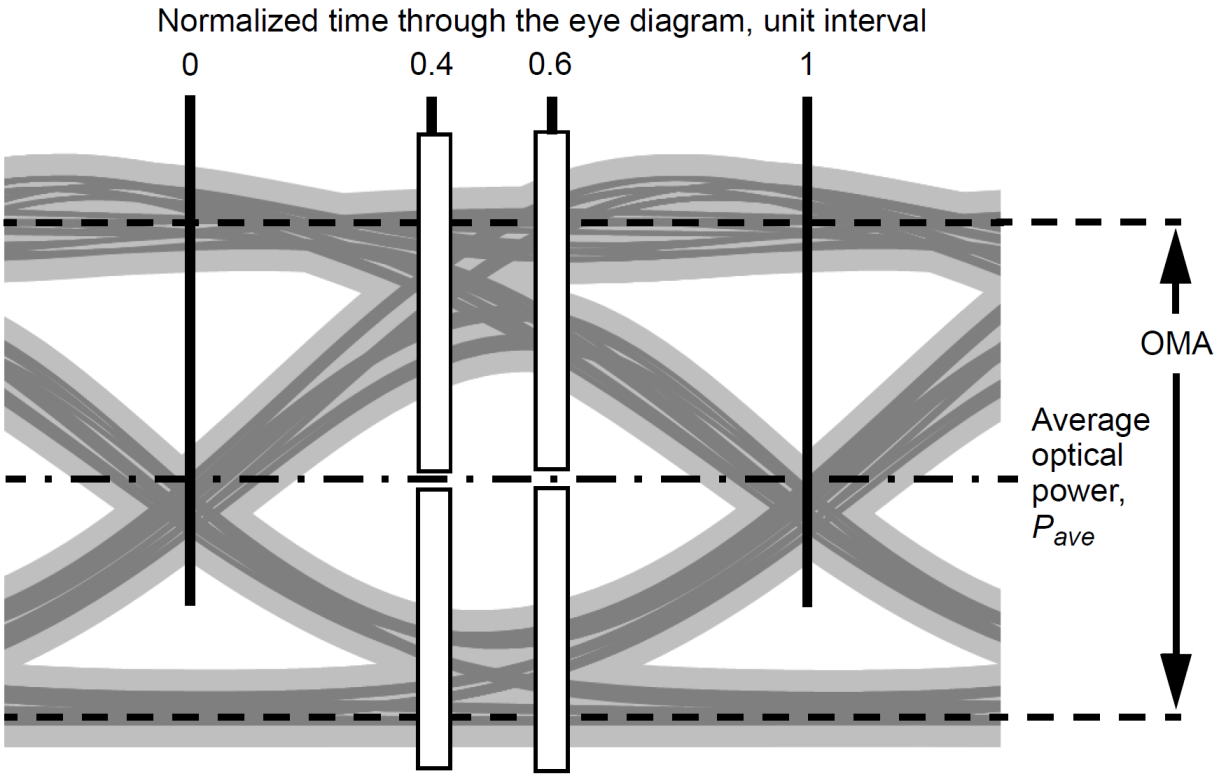


Figure 2-47. *Illustration of the TDEC measurement.*

OMA is measured according to Section **2.9.4.3**.

The standard deviation of the noise of the O/E and oscilloscope combination, *S*, is determined with no optical input signal and the same settings as used to capture the histograms described below.

The average optical power (P_{ave}), the crossing points of the eye diagram, and the four vertical histograms used to calculate TDEC, are all measured using the same selected test pattern from those identified for TDEC in **Table 2-42**. The 0 UI and 1 UI crossing points are determined by the average of the eye diagram crossing times, as measured at P_{ave}, as illustrated in **Figure 2-47**.

Four vertical histograms are measured through the eye diagram, centered at 0.4 UI and 0.6 UI, and above and below P_{ave}, as illustrated in **Figure 2-47**.

Each histogram window has a width of 0.04 UI. Each histogram window has an inner height boundary, which is set close to P_{ave} (so that no further samples would be captured by moving it closer to P_{ave}), and an outer height boundary, which is set beyond the outer-most samples of the eye diagram (so that no further samples would be captured by increasing the outer histogram window boundary).

The distributions of the two histograms on the left are each multiplied by *Q* functions, which represent an estimate probability of errors caused by each part of the distribution for the greatest tolerable noise that could be added by an optical channel and a receiver. The resulting distributions are integrated, and each integral is divided by the integral of the distribution it was derived from, giving two-bit error probabilities. The *Q* function uses a standard deviation, σ_L, chosen so that the average of these two-bit error probabilities is 5×10^{-5}. Similarly, for the two histograms on the right, a standard deviation, σ_R, is found.

Q(x) is the area under a normal curve for values larger than *x* (the tail probability, related to the "complementary error function"), as shown in **Equation 2-8**:

$$\mathrm{Q(x)} = \int_x^{\infty} \frac{e^{\left(\frac{-z^2}{2}\right)}}{\sqrt{2\pi}} d_z$$

Equation 2-8

where

x is $(y–P_{ave})/\sigma_G$ or $(P_{ave}–y)/\sigma_G$, as in **Equation 2-9**

This procedure finds a value of σ_G such that **Equation 2-9** is satisfied:

$$\frac{1}{2}\left(\frac{\int fu(y)Q\left(\frac{y-P_{ave}}{\sigma_G}\right)dy}{\int f_u(y)dy}\right)+\frac{1}{2}\left(\frac{\int fl(y)Q\left(\frac{P_{ave}-y}{\sigma_G}\right)dy}{\int fl(y)dy}\right)=5\times10^{-5}$$

Equation 2-9

where

fu(y), fl(y) are the upper and lower distributions,

σ_G is the left or right standard deviation, σ_L or σ_R,

The lesser of σ_L and σ_R is *N*.

The noise, *R*, that could be added by a receiver is given by **Equation 2-10**.

$$R = (1 - M_1)\sqrt{N^2 + S^2 - M_2^2}$$

Equation 2-10

where

M_1, M_2 defined in **Equation 2-11** and **Equation 2-12**, account for mode partition noise and modal noise that could be added by the optical channel,

S is the standard deviation of the noise of the O/E and oscilloscope combination

$$M_1 = 0.04$$

Equation 2-11

$$M_2 = 0.0175P_{ave}$$

Equation 2-12

where

P_{ave} is the average optical power of the eye diagram.

TDEC is given by **Equation 2-13**.

$$TDEC = 10log_{10}\left(\frac{OMA}{2} \times \frac{1}{3.8906R}\right)$$

Equation 2-13

where

OMA is the optical modulation amplitude as defined in Section **2.9.4.3**.

The factor 3.8906 is chosen for consistency with the BER of 5×10^{-5} given in Section **2.9**.

The method described in Sections **2.9.4.5** and **2.9.4.6** is the reference measurement method.

2.9.4.7 Extinction ratio

The extinction ratio of each lane shall be within the limits given in **Table 2-40** if measured using the methods defined in IEC 61280-2-2. The extinction ratio is measured using one of the test patterns specified for extinction ratio in **Table 2-42**.

NOTE: Extinction ratio and OMA are defined with different test patterns (see **Table 2-42**).

2.9.4.8 Transmitter optical waveform (transmit eye)

The required shape characteristics of the optical transmitter pulse are specified in the form of a mask of the transmitter eye diagram as shown in **Figure 2-22** with the transmitter eye mask coordinates and hit ratio as in **Table 2-39**. The transmitter optical waveform of a port transmitting the test pattern specified in **Table 2-42** shall meet specifications in compliance with the methods specified in Section **2.6.4.8** with the following exceptions:

- The clock recovery unit's high-frequency corner bandwidth is 10MHz.
- The filter nominal reference frequency *fr* is 19.34GHz and the filter tolerances are as specified for STM-64 in ITU-T G.691.

2.9.4.9 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in **Table 2-39** if measured using the method defined by Sections **2.9.4.10** and **2.9.4.14**, with the conformance test signal at TP3 as described in Section **2.9.4.11**.

Stressed receiver sensitivity is defined with all transmit and receive lanes in operation. Pattern 3 or Pattern 5, or a valid 100GBASE-SR4 signal is sent from the transmit section of the PMD under test. The signal transmitted is asynchronous to the received signal. The interface BER of the PMD receiver is the average of the BER of all receive lanes while stressed and at the specified receive *OMA*.

2.9.4.10 Stressed receiver conformance test block diagram

A block diagram for the receiver conformance test is shown in **Figure 2-48**. The patterns used for the received conformance signal are specified in **Table 2-42**. The optical test signal is conditioned (stressed) using the stressed receiver methodology specified in Section **2.9.4.11** and has sinusoidal jitter applied as defined in Section **2.9.4.14**. A suitable test set is necessary to verify that the signal used to test the receiver has the appropriate characteristics. The fourth-order Bessel-Thomson filter has a 3dB bandwidth of approximately 19GHz. The low-pass filter is used to create *ISI*. The combination of the low-pass filter and the E/O converter should have a frequency response that results in the level of Stressed Eye Closure (SEC) before the sinusoidal and Gaussian noise terms are added, as described in Section **2.9.4.11**.

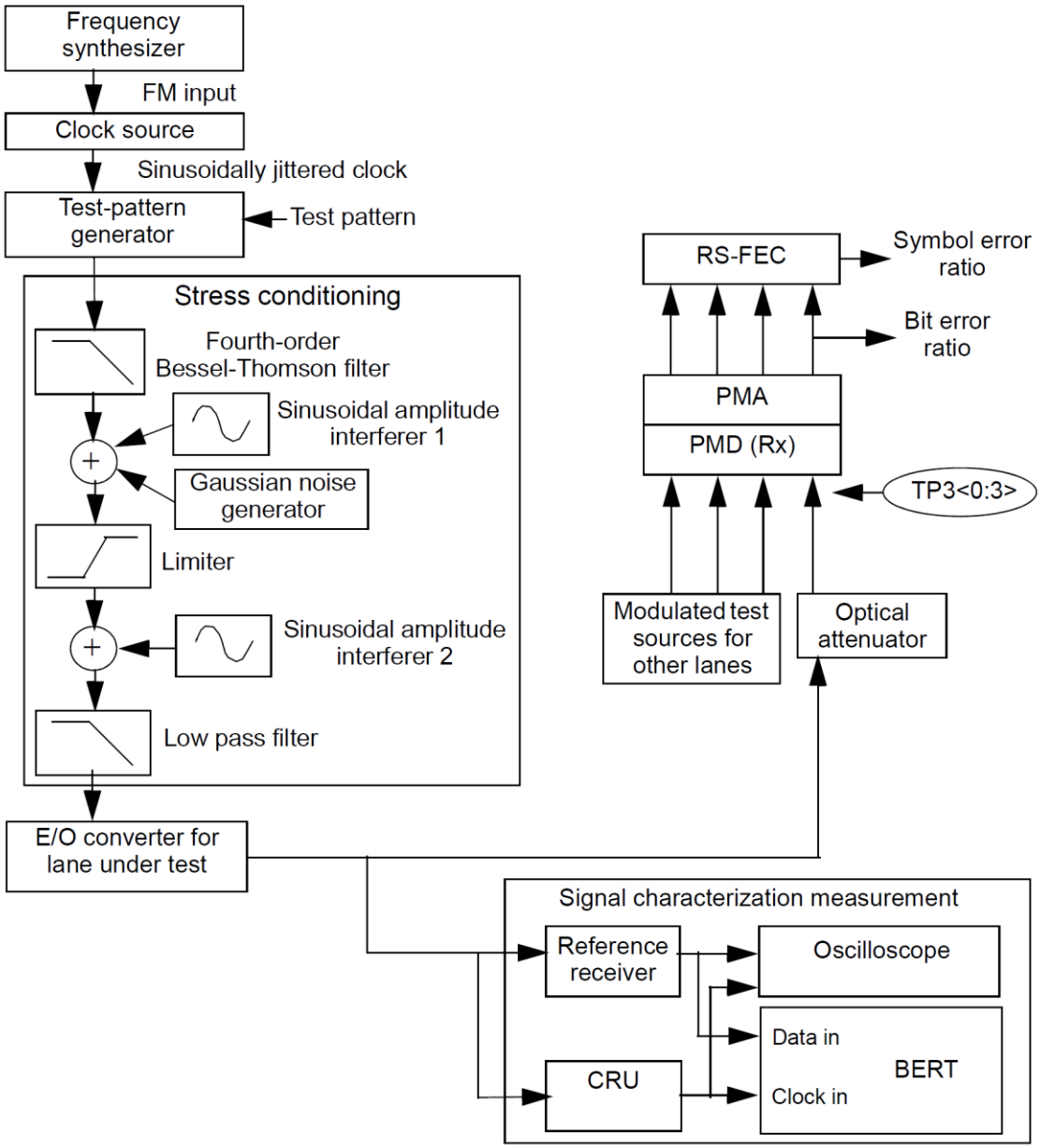


Figure 2-48. *Stressed receiver conformance test block diagram.*

The sinusoidal amplitude interferer 1 causes jitter, which is intended to emulate instantaneous bit shrinkage that happens with Data Depend Jitter (DDJ). This type of jitter cannot be created by simple phase modulation. The sinusoidal amplitude interferer 2 causes additional eye closure, but in conjunction with the finite edge rates from the limiter, which also causes jitter. The sinusoidally jittered clock represents other types of jitter and also verifies that the receiver under test can track low-frequency jitter. The sinusoidal amplitude interferers is set at any frequency between 100MHz and 2GHz, although provision should be taken to avoid harmonic relationships between the sinusoidal interferers, the sinusoidal jitter, the signaling rate,

and also the pattern repetition rate. The Gaussian noise generator, the amplitude of the sinusoidal interferers, and the low-pass filter are adjusted so that the SEC and stressed eye J2 Jitter specifications given in **Table 2-40** are met simultaneously, while the maximum stressed eye J4 Jitter defined in **Table 2-40** is not exceeded, and also while passing the stressed receiver eye mask defined in **Table 2-40** in compliance with the methods defined in Section **2.9.4.8** (the random noise effects such as RIN, or random clock jitter, do not need to be minimized).

The stressed receiver conformance test signal verification is specified in Section **2.9.4.13**.

Stressed receiver sensitivity is specified with all transmit and receive lanes in operation. Each receive lane is tested in turn, while all aggressor receive lanes are operated as defined in **Table 2-40**. Pattern 3, Pattern 5, or any valid 100GBASE–SR4 signal is sent from the transmit section of the receiver under test. The transmitted signal is asynchronous to the received signal. If Pattern 3 is used with a common clock for the transmit or receive lanes not under test, there is a delay of at least 31 UI between the PRBS31 patterns generated on one lane and any other lane.

The relevant BER of 100GBASE-SR4 is the interface BER at the PMD service interface. The interface BER is the average of the four BER of the receive lanes when stressed. If the RS-FEC sublayer is present, the measurement of the lane symbol error ratio can be made at its input. The lane BER can be assumed to be one tenth of the lane symbol error ratio. If each lane is stressed in turn, the PMD interface BER is the average of the BERs of all the lanes when stressed.

2.9.4.11 Stressed receiver conformance test signal characteristics and calibration

The conformance test signal must validate that each lane of the PMD receiver complies with the BER requirements of the near worst-case waveforms at TP3.

The primary parameters of the stressed receiver conformance test signal are its SEC, the stressed eye J2 Jitter, and the stressed eye J4 Jitter. The SEC of the stressed receiver conformance test signal is measured as specified in Section **2.9.4.4**, apart the combination of the O/E and the oscilloscope used to measure the waveform has a fourth-order Bessel-Thomson filter response with a bandwidth of 19.34GHz, and the values of M1 and M2 set to zero in **Equation 2-10**. Stressed eye J2 Jitter and stressed eye J4 Jitter are defined in Section **2.9.4.12**.

An example stressed receiver conformance test setup is shown in **Figure 2-48**. However, alternative test setups may be used to generate equivalent stress conditions.

The following steps provide a method for setting up and calibrating a stressed receiver conformance test signal when using a stressed receiver conformance test setup as shown in **Figure 2-48**:

1. Set the signaling rate of the test pattern generator to meet the requirements in **Table 2-40**.
2. With the sinusoidal jitter, sinusoidal interferers, and the Gaussian noise generator turned off, set the extinction ratio of the E/O to approximately the minimum specified in **Table 2-39**.
3. The required values of SEC and J2 Jitter, and the maximum value of J4 Jitter of the stressed receiver conformance test signal are given in **Table 2-40**. A valid stressed receiver conformance test signal may have a lower value of J4 jitter than the maximum specified in **Table 2-40**, provided to meet the required values of SEC and J2 Jitter.

   With the sinusoidal jitter, the sinusoidal interferer 1, the sinusoidal interferer 2, and the Gaussian noise generator turned off, a SEC of at least 2.5dB should be created by selecting the appropriate bandwidth for both the low-pass filter and the E/O converter. Any remaining SEC must be created with a combination of sinusoidal jitter, sinusoidal interference, and Gaussian noise.

   Sinusoidal jitter is added as per **Table 2-44**. The sinusoidal jitter frequency should be between 50MHz and 10 times Loop Bandwidth (LB), when calibrating the conformance signal as per **Table 2-44**. Sinusoidal jitter amplitude must be calibrated by measuring the jitter on the oscilloscope, while transmitting the square wave pattern, and using a clean clock in place of the CRU to trigger the oscilloscope.

   Iterate the adjustments of sinusoidal interferers and Gaussian noise generator and extinction ratio until the values of SEC and stressed eye J2 Jitter are met, while also meeting the following conditions: the maximum value of stressed eye J4 Jitter, as per **Table 2-40**, is not exceeded; the extinction ratio is approximately the minimum as per **Table 2-39**, and sinusoidal jitter is as per **Table 2-44**.

Each receiver lane is conformance tested in turn. The source for the lane under test is adjusted to supply a signal at the input of the receiver under test at the “Stressed receiver sensitivity (OMA), each lane (max)” as per **Table 2-40**, and the test sources for the other lanes are set to the “OMA of each aggressor lane” as per **Table 2-40**.

2.9.4.12 J2 and J4 Jitter

Jitter J2 is defined as the time interval at the average optical power level, which includes all but 10^{-2} of the jitter distribution, as the time interval from the 0.5th to the 99.5th percentile of the jitter histogram. Jitter J2 is defined using a clock recovery unit as per Section **2.9.4.8**. If measured using an oscilloscope, the histogram should include at least 10 000 hits and should be about 1% of the signal amplitude. If measured by plotting BER against the decision time, J2 is the time interval between the two points with a BER of 2.5×10^{-3}.

Jitter J4 is defined as the time interval at the average optical power level that includes all but 10^{-4} of the jitter distribution. Jitter J4 is defined using a clock recovery unit as per Section **2.9.4.8.** If measured using an oscilloscope, the histogram should include at least 1 000 000 hits and should be about 1% of the signal amplitude. If measured by plotting BER against the decision time, J4 is the time interval between the two points with a BER of 2.5×10^{-5}.

2.9.4.13 Stressed receiver conformance test signal verification

The stressed receiver conformance test signal can be verified using an optical reference receiver with an ideal fourth-order Bessel-Thomson response with a reference frequency *fr* of 19.34GHz. The use of the ITU-T G.691 tolerance filters may significantly degrade this calibration. The clock output from the clock source shown in **Figure 2-48** is modulated with the sinusoidal jitter. To use an oscilloscope to calibrate the final stressed eye J2 Jitter and stressed eye J4 Jitter that includes the sinusoidal jitter component, a clock recovery unit (CRU of **Figure 2-48**) is required.

when characterizing the test signal care should be taken because an excessive noise/jitter in the measurement system would result in an input signal that does not fully stress the receiver under test. Running the receiver tolerance test with a signal that is under-stressed may result in the deployment of non-compliant receivers. Care should be taken to minimize the noise/jitter introduced by the reference O/E, filters, and BERT and/or to correct for this noise. Although the details of a BER scan measurement and test equipment are beyond the scope of this standard, it is

recommended that the implementer fully characterize the test equipment and apply appropriate guard bands to ensure that the stressed receiver conformance input signal meets the stress and sinusoidal jitter as per Sections **2.9.4.11** and **2.9.4.14**.

2.9.4.14 Sinusoidal jitter for receiver conformance test

The sinusoidal jitter is used for the testing of the receiver jitter tolerance. The amplitude of the applied sinusoidal jitter depends on the frequency as per **Table 2-44**.

Frequency range	Sinusoidal jitter peak-to-peak (UI)
f < 100kHz	Not specified
100kHz < f ≤ 10MHz	$5 \times 10^5/f$
10MHz < f ≤ 10 LB [a]	0.05
[a] the upper frequency bound for added sinusoidal jitter should be at least 10 times the LB of the receiver being tested.	

Table 2-44. *Applied sinusoidal jitter.*

2.9.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

100GBASE-SR4 optical transceivers shall conform to Hazard Level 1M laser requirements in compliance with IEC 60825-1 and IEC 60825-2, under any condition of operation.

Laser safety standards and regulations require the manufacturer of a laser product shall provide information about the product's laser, safety features, labeling, use, maintenance, and service.

It is recommended that proper installation practices, as defined by the applicable local codes and regulations, shall be followed in every instance.

It is also recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of this Section shall be met.

A system integrating a 100GBASE-SR4 PMD shall comply with the applicable local and national codes for the limitation of electromagnetic interference.

It is recommended that each PHY (and supporting documentation) be labeled in a manner visible to the user, with at least the applicable safety warnings and the applicable port type designation (e.g., 100GBASE-SR4).

2.9.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 2-49**.

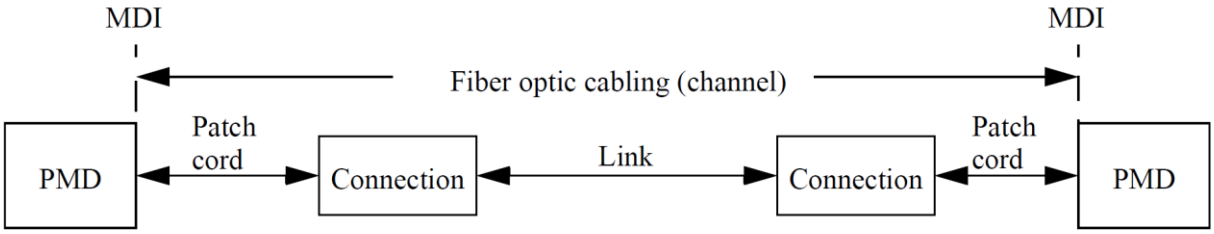


Figure 2-49. *Fiber optic cabling model*

The channel insertion loss is given in **Table 2-45**. A channel can contain additional connectors and splices as long as the optical characteristics of the channel (such as attenuation, modal dispersion, reflections and losses of all connectors and splices) in compliance with the appropriate specifications. Insertion loss measurements of installed fiber cables shall be in compliance with IEC 61280-4-1. As OM4 optical fiber complies with the requirements for OM3, a channel compliant to the OM3 column may use OM4 optical fiber, or a combination of OM3 and OM4. The fiber optic cabling model, known as channel, is the same as a simplex fiber optic link segment. The term channel is used for consistency with the standards of the generic cabling.

Description	OM3	OM4	Unit
Operating distance (max)	70	100	m
Cabling Skew (max)	79		ns
Cabling Skew Variation [a] (max)	2.4		ns
Channel insertion loss [b] (max)	1.8	1.9	dB
Channel insertion loss (min)	0		dB
[a] An additional Skew Variation of 400ps could be caused by the wavelength changes, which are attributable to the transmitter and not the channel. [b] These values of the channel insertion loss include the cable loss plus 1.5dB allocated for each connection and splice loss, over the wavelength range from 840nm to 860nm.			

Table 2-45. *Fiber optic cabling (channel) characteristics for 100GBASE-SR4.*

2.9.7 Characteristics of the fiber optic cabling (channel)

The 100GBASE-SR4 fiber optic cabling shall comply with the specifications defined in **Table 2-45**. The fiber optic channel consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

The fiber contained within the 100GBASE-SR4 fiber optic cabling shall comply with the specifications and parameters of **Table 2-46**. A variety of multimode cable types

can satisfy these requirements, and the resulting channel shall also comply with the specifications of **Table 2-45**.

Description	OM3 [a]	OM4 [b]	Unit
Nominal core diameter	50		μm
Nominal fiber specification wavelength	850		nm
Effective modal bandwidth (min) [c]	2000	4700	MHz·km
Cabled optical fiber attenuation (max)	3.5		dB/km
Zero dispersion wavelength (λ_0)	$1295 \le \lambda_0 \le 1340$		nm
Chromatic dispersion slope (max) (S_0)	0.105 for $1295 \le \lambda_0 \le 1310$ and $0.000375 \times (1590 - \lambda_0)$ for $1310 \le \lambda_0 \le 1340$		ps/nm^2 km
[a] IEC 60793-2-10 type A1a.2 [b] IEC 60793-2-10 type A1a.3 [c] When measured with the launch conditions specified in **Table 2-39**			

Table 2-46. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 2-49**, consists of a mated pair of optical connectors.

The maximum link distance is based on the allocation of 1.5dB total connection and splice loss. For example, this allocation supports three connections with an average insertion loss per connection of 0.5dB. Connections with lower loss characteristics may be used provided the requirements of **Table 2-45** are met. However, the total loss of a single connection shall not exceed 0.75dB.

The maximum discrete reflectance shall be less than –20dB.

2.9.7.1 Medium Dependent Interface (MDI)

The 100GBASE-SR4 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in **Figure 2-49**). The 100GBASE-SR4 PMD is coupled to the fiber optic cabling through one connector plug into the MDI optical connector as shown in **Figure 2-50**. Example constructions of the MDI include the following:

a. PMD with a connectorized fiber pigtail plugged into an adapter.
b. PMD connector.

2.9.7.2 Optical lane assignments

The four transmit and four receive optical lanes of 100GBASE-SR4 shall occupy the positions depicted in **Figure 2-50** when looking into the MDI connector with the keyway feature on top. The interface contains eight active lanes distributed within twelve total

positions. The transmit optical lanes allocate the left-most four positions. The receive optical lanes allocate the right-most four positions. The four center positions are unused.

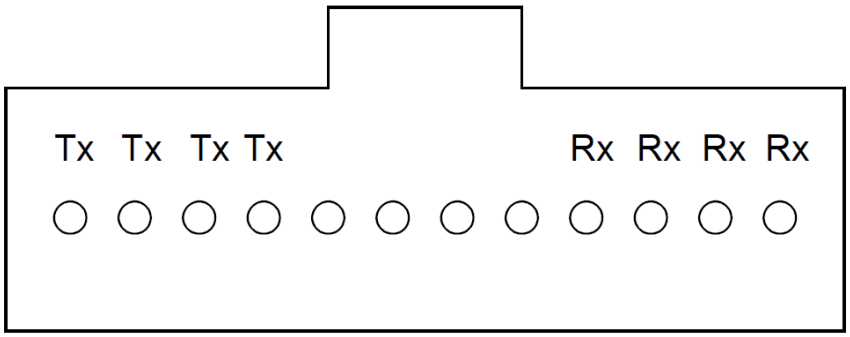


Figure 2-50. *100GBASE-SR4 optical lane assignments.*

2.9.7.3 Medium Dependent Interface (MDI) requirements

The MDI adapter and connector shall meet the dimensional specifications for interface 7-1-3: *MPO adapter interface – opposed keyway configuration, or interface 7-1-10: MPO active device connector, flat interface*, as defined in IEC 61754-7-1. The plug terminated with the optical fiber cabling shall comply with the dimensional specifications of interface 7-1-4: *MPO female plug connector, flat interface for 2 to 12 fibres*, as specified in IEC 61754-7-1. The MDI shall be optically mated with the plug on the optical fiber cabling. **Figure 2-51** shows an MPO female plug connector with flat interface, and an MDI.

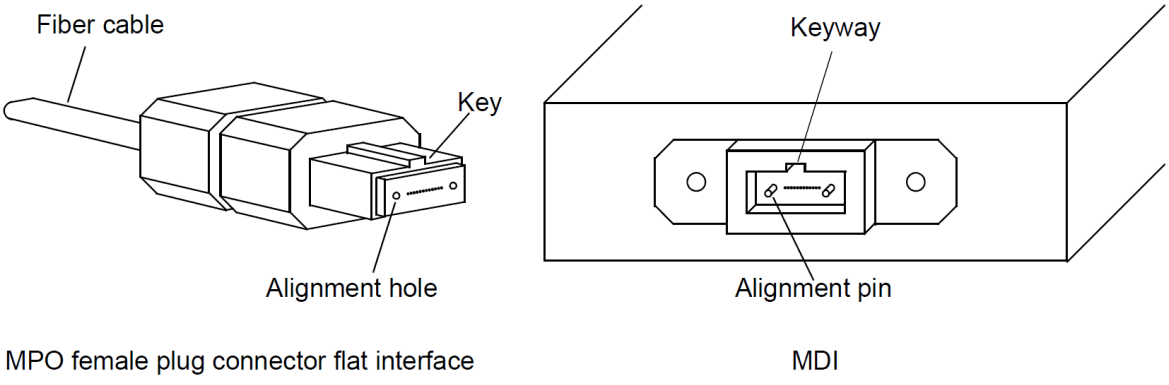


Figure 2-51. *MPO female plug with flat interface and MDI*

The MDI shall meet the interface performance specifications of the IEC 61753-1 and IEC 61753-22-2.

NOTE: Transmitter compliance testing is performed at TP2 as per Section **2.9.1**, and not at the MDI.

3 Architectural Perspectives of 200Gb/s and 400Gb/s Ethernet

200 Gigabit and 400 Gigabit Ethernet couples the IEEE 802.3 MAC to a family of 200Gb/s and 400Gb/s PHYs. The relationships among 200 Gigabit and 400 Gigabit Ethernet, the IEEE 802.3 MAC, and the ISO OSI reference model are shown in **Figure 3-1**.

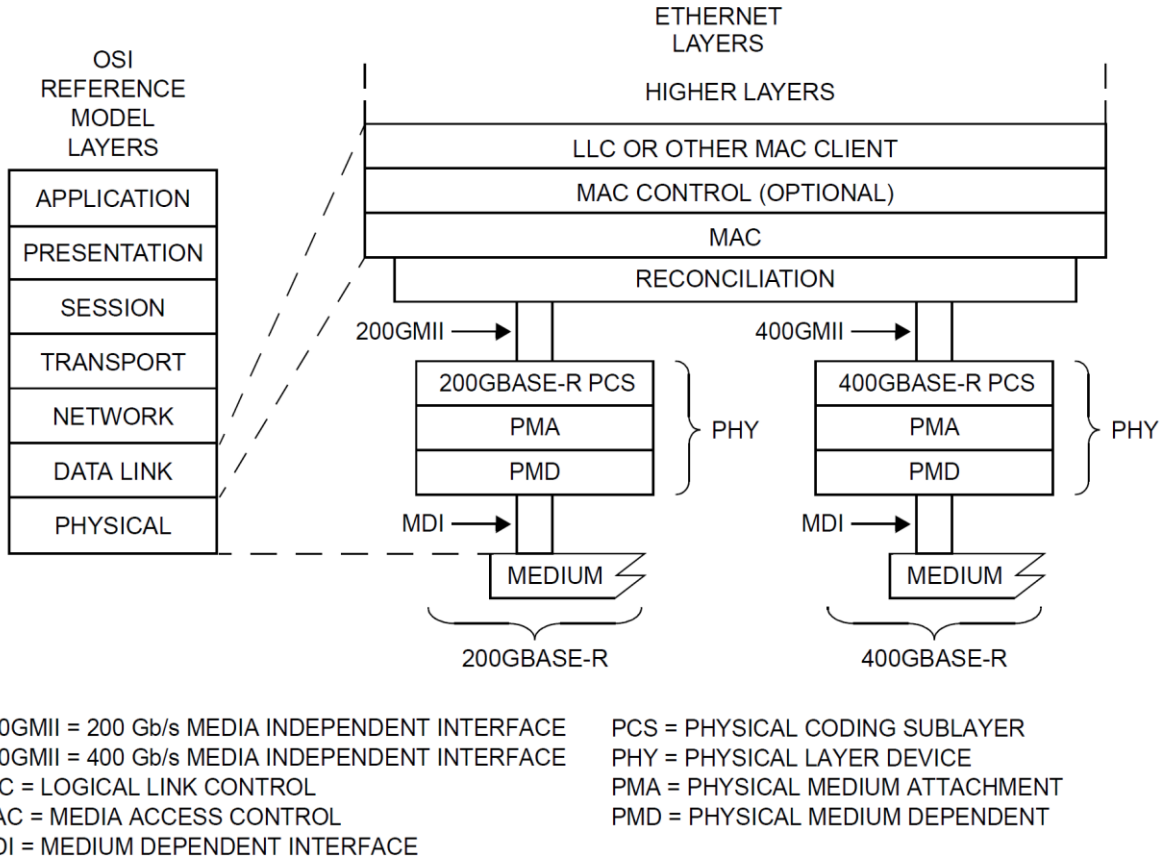


Figure 3-1. *Architecture of 200Gb/s and 400Gb/s Ethernet.*

The terminology used in 200Gb/s and 400Gb/s PHY is described as follows:

- The alpha-numeric prefix 200GBASE in the port type (e.g., 200GBASE-R) represents a family of PHY devices operating at a speed of 200 Gb/s.
- The alpha-numeric prefix 400GBASE in the port type (e.g., 400GBASE-R) represents a family of PHY devices operating at a speed of 400 Gb/s.

200GBASE-R represents a family of PHY devices using the PCS for 200 Gb/s operation over multiple PCS lanes. PHY devices listed in **Table 3-1** are defined for operation at 200 Gb/s.

Name	Description
200GBASE-DR4	200 Gb/s PHY using 200GBASE-R encoding over four lanes of single-mode fiber, with reach up to at least 500m
200GBASE-FR4	200 Gb/s PHY using 200GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 2km
200GBASE-LR4	200 Gb/s PHY using 200GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 10km

Table 3-1. *200Gb/s PHYs*

400GBASE-R represents a family of PHY devices using the PCS sublayer for 400 Gb/s operation over multiple PCS lanes. PHY devices listed in **Table 3-2** are defined for operation at 400 Gb/s.

Name	Description
400GBASE-SR16	400 Gb/s PHY using 400GBASE-R encoding over sixteen lanes of multimode fiber, with reach up to at least 100m
400GBASE-DR4	400 Gb/s PHY using 400GBASE-R encoding over four lanes of single-mode fiber, with reach up to at least 500m
400GBASE-FR8	400 Gb/s PHY using 400GBASE-R encoding over eight WDM lanes on single-mode fiber, with reach up to at least 2km
400GBASE-LR8	400 Gb/s PHY using 400GBASE-R encoding over eight WDM lanes on single-mode fiber, with reach up to at least 10km

Table 3-2. *400Gb/s PHYs.*

3.1 Media Access Control (MAC)

3.1.1 MAC Service Specifications

The MAC control sublayer is a client of the CSMA/CD MAC. **Figure 3-2** depicts the architectural positioning of the MAC Control sublayer with respect to the CSMA/CD MAC and the MAC Control client. MAC Control clients may include the Bridge Relay Entity (BRE), Logical Link Control (LLC), or other applications.

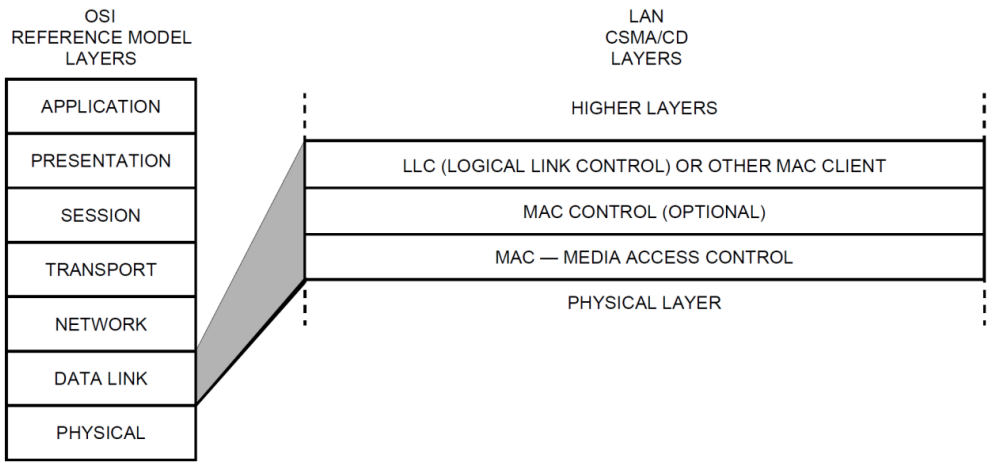


Figure 3-2. *MAC control sublayer.*

The MAC sublayer provides services to the client of the MAC. MAC clients may include the LLC sublayer, BRE, or other users of ISO/IEC LAN International Standard MAC services described in **Figure 3-3**.

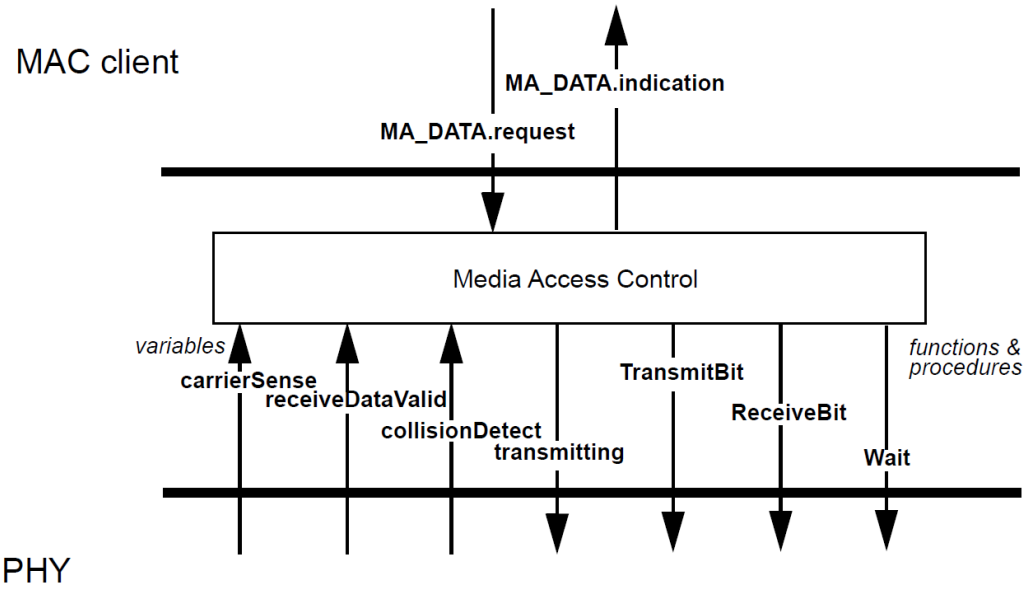


Figure 3-3. *MAC services.*

The services provided by the MAC sublayer allow the local MAC client entity to exchange LLC data units with peer LLC sublayer entities.

3.1.2 MAC Frame and Packet Specifications

In Ethernet's history, many capabilities have been added to allow data link layer (layer 2) protocol encapsulations within the MAC Client Data field. So, there are now more than one type of MAC frame.

The frame format specified in IEEE Std. 802.3 series includes the following three types of MAC frames:

a. A basic frame.
b. A Q-tagged frame.
c. An envelope frame.

All three frame types use the same Ethernet frame format.

Figure 3-4 shows the fields of a packet:

i. the Preamble,
ii. the Start Frame Delimiter (SFD),
iii. the addresses of the MAC frame's destination and source,
iv. a length or type field to indicate the length or protocol type of the following field that contains the MAC client data,
v. a field that contains padding if required, and
vi. the Frame Check Sequence (FCS) field containing a cyclic redundancy check value to detect errors in a received MAC frame.

Of these fields, all are of fixed size except for the MAC Client Data, Pad and Extension fields, which may contain an integer number of octets between the minimum and the maximum values, which are determined by the specific implementation of the MAC.

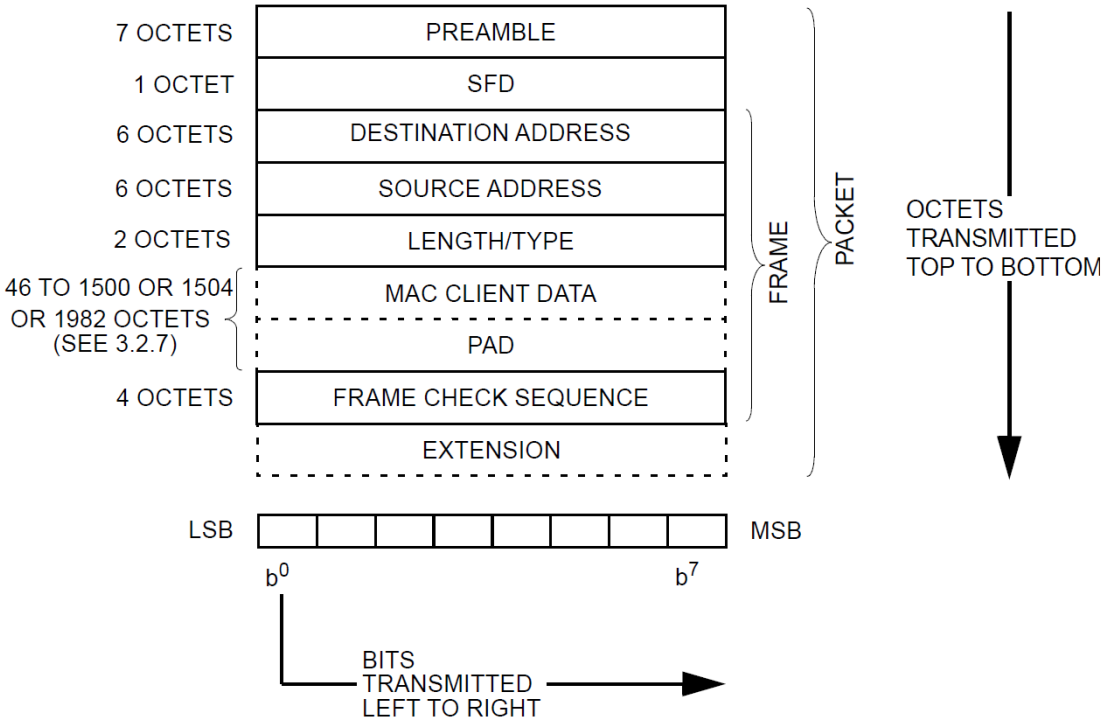


Figure 3-4. *MAC packet format.*

3.1.3 Delay constraints

Predictable operation of the MAC Control PAUSE operation demands that there be an upper bound on the propagation delays through the network. This implies that the MAC, the MAC Control sublayer, and the PHY implementers shall conform to certain delay maxima, and that network planners and administrators shall conform to constraints, regarding the cable topology and concatenation of the devices.

Table 3-3 and **Table 3-4** contain the values of maximum sublayer delay (the sum of transmit and receive delays at the one end of the link) in bit times and the *pause_quanta* for 200 Gigabit and 400 Gigabit Ethernet, respectively. If a PHY contains an Auto-Negotiation sublayer, then the delay of the Auto-Negotiation sublayer is included within the delay of the PMD and the medium.

Sublayer	Maximum (bit time) [a]	Maximum (pause_quanta) [b]	Maximum (ns)	Notes
200G MAC, RS, and MAC Control	49,152	96	245.76	
200GBASE-R PCS or 200GXS [c]	160,256	313	801.28	
200GBASE-R PMA	18,432	36	92.16	
200GBASE-DR4 PMD	4,096	8	20.48	Includes 2m of fiber.
200GBASE-FR4 PMD	4,096	8	20.48	Includes 2m of fiber.
200GBASE-LR4 PMD	4,096	8	20.48	Includes 2m of fiber.

[a] For 200GBASE-R, 1 Bit Time (BT) is equal to 5ps.

[b] For 200GBASE-R, 1 *pause_quantum* is equal to 2.56ns.

[c] If an implementation includes the 200GMII extender, the delay associated with the 200GMII extender includes two 200GXS sublayers.

Table 3-3. *Sublayer delay constraints (200GBASE).*

Sublayer	Maximum (bit time) [a]	Maximum (pause_quanta) [b]	Maximum (ns)	Notes
400G MAC, RS, and MAC Control	98,304	192	245.76	
400GBASE-R PCS or 400GXS [c]	320,000	625	800	

400GBASE-R PMA	36,864	72	92.16	
400GBASE-SR16 PMD	8,192	16	20.48	Includes 2m of fiber.
400GBASE-DR4 PMD	8,192	16	20.48	Includes 2m of fiber.
400GBASE-FR8 PMD	8,192	16	20.48	Includes 2m of fiber.
400GBASE-LR8 PMD	8,192	16	20.48	Includes 2m of fiber.
[a] For 400GBASE-R, 1 BT is equal to 2.5ps. [b] For 400GBASE-R, 1 *pause_quantum* is equal to 1.28ns. [c] If an implementation includes the 400GMII extender, the delay associated with the 400GMII extender includes two 400GXS sublayers.				

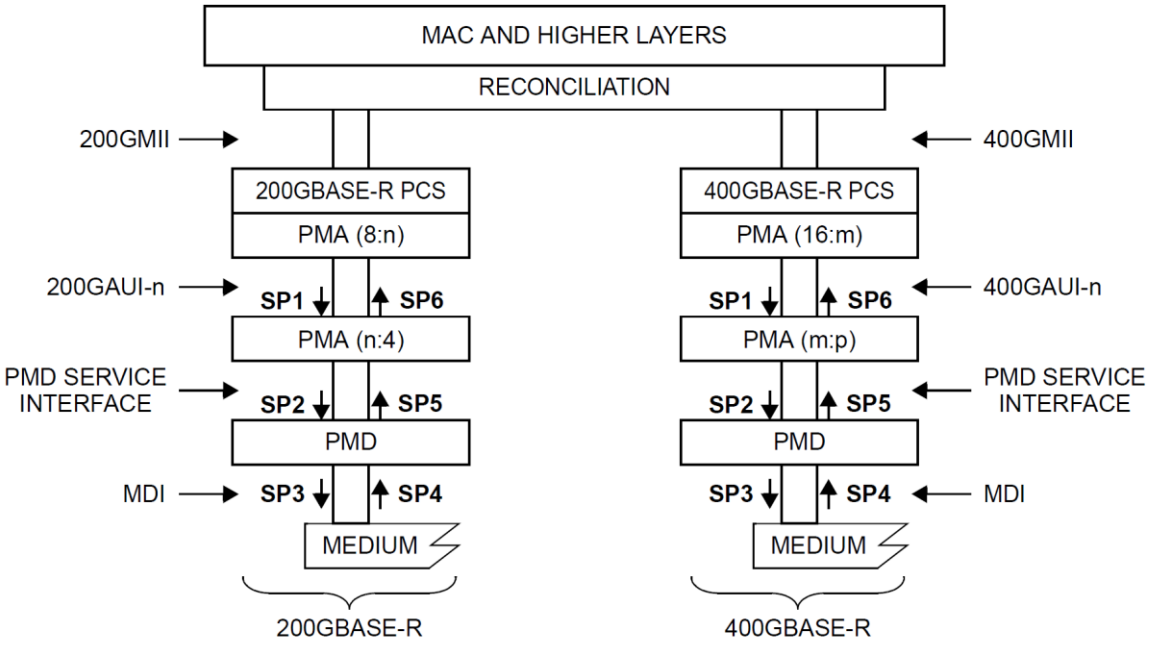
Table 3-4. *Sublayer delay constraints (400GBASE).*

3.1.4 Skew constraints

Skew (or relative delay) can be introduced between lanes by both active and passive elements of a 200GBASE-R or 400GBASE-R link. Skew is defined as the difference between the times of the earliest and the latest PCS lane for the one to zero transition of the alignment marker sync bits. The PCS deskew function compensates for all lane-to-lane Skew observed at the receiver. The Skew between the lanes must be kept within limits as shown in **Table 3-5** so that the transmitted information on the lanes can be reassembled by the receive PCS.

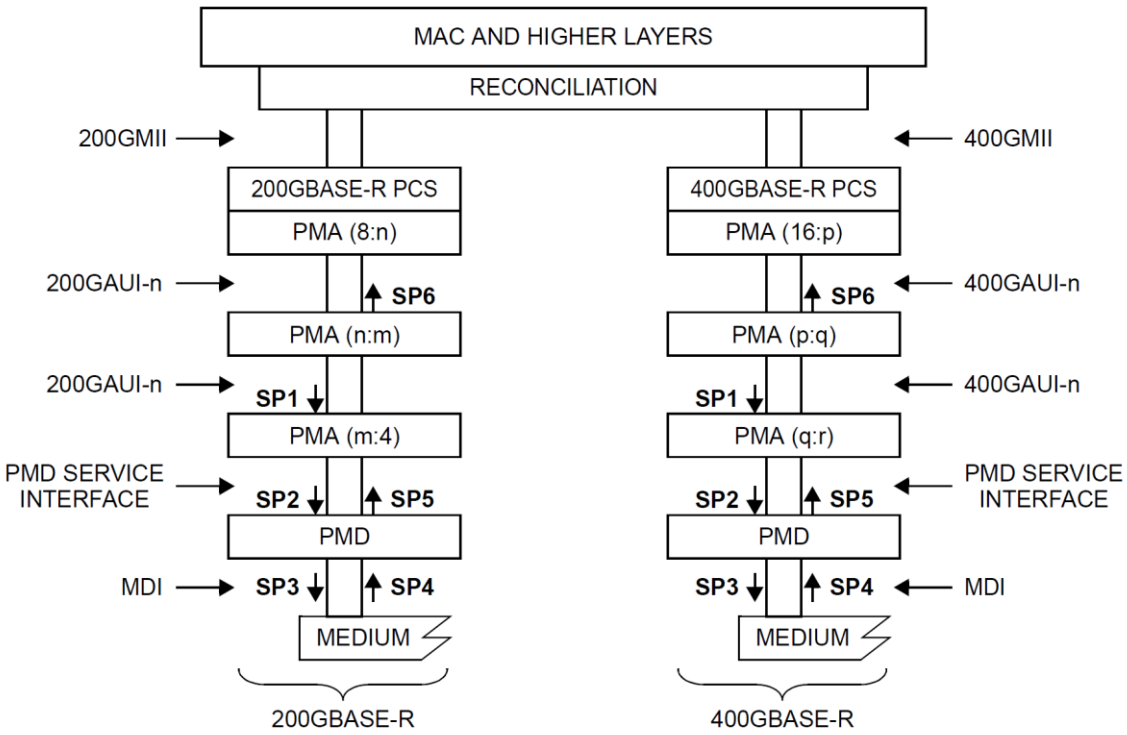
Skew Variation is introduced due to variations in electrical, thermal, or environmental characteristics. Skew Variation is the change in Skew between any PCS lane and any other PCS lane over the entire time that the link is in operation. From the time the link is brought up, Skew Variation is limited so that each PCS lane always traverses the same lane between any pair of adjacent sublayers while the link remains in operation.

The maximum Skew and Skew Variation at physically instantiated interfaces is specified at the Skew points SP1, SP2, and SP3 for the transmit direction and SP4, SP5, and SP6 for the receive direction as illustrated in **Figure 3-5** (single 200GAUI-n or 400GAUI-n interface) and **Figure 3-6** (multiple 200GAUI-n or 400GAUI-n interfaces).



200GAUI-n = 200 Gb/s ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI-n = 400 Gb/s ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
n = 8 or 4
m = 16 or 8
p = 16, 8, or 4

Figure 3-5. *200GBASE-R and 400GBASE-R Skew points for single 200GAUI-n or 400GAUI-n.*



200GAUI-n = 200 Gb/s ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI-n = 400 Gb/s ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
n = 8 or 4
m = 8 or 4
p = 16 or 8
q = 16 or 8
r = 16, 8, or 4

Figure 3-6. *200GBASE-R and 400GBASE-R Skew points for multiple 200GAUI-n or 400GAUI-n.*

In the transmit direction, the Skew points are defined in the following locations:

- SP1 on the 200GAUI-n/400GAUI-n interface, at the input of the PMA closest to the PMD,
- SP2 on the PMD service interface, at the input of the PMD,
- SP3 at the output of the PMD, at the MDI.

In the receive direction, the Skew points are defined in the following locations:

- SP4 at the MDI, at the input of the PMD,
- SP5 on the PMD service interface, at the output of the PMD,
- SP6 on the 200GAUI-n/400GAUI-n interface, at the output of the PMA closest to the 200GBASE-R/400GBASE-R PCS or Data Terminal Equipment (DTE) 200GXS/400GXS.

The allowable limits for Skew are shown in **Table 3-5** and the allowable limits for Skew Variation are shown in **Table 3-6**.

Skew points	Maximum Skew (ns) [a]	Maximum Skew for 200GBASE-R or 400GBASE-R PCS lane (UI) [b]
SP1	29	≈ 770
SP2	43	≈ 1142
SP3	54	≈ 1434
SP4	134	≈ 3559
SP5	145	≈ 3852
SP6	160	≈ 4250
At PCS receive	180	≈ 4781

[a] The Skew limit includes 1ns allowance for PCB traces that are associated with the Skew points.

[b] The symbol ≈ indicates approximate equivalent of maximum Skew in UI based on 1 UI equals 37.64706ps at PCS lane signaling rate of 26.5625GBd.

Table 3-5. *Summary of Skew constraints.*

Skew points	Maximum Skew Variation (ns)	Maximum Skew Variation for 26.5625GBd PMD lane (UI) [a]	Maximum Skew Variation for 53.125GBd PMD lane (UI) [b]
SP1	0.2	≈ 5	N/A
SP2	0.4	≈ 11	≈ 21
SP3	0.6	≈ 16	≈ 32
SP4	3.4	≈ 90	≈ 181
SP5	3.6	≈ 96	≈ 191
SP6	3.8	≈ 101	N/A
At PCS receive	4	≈ 106	N/A

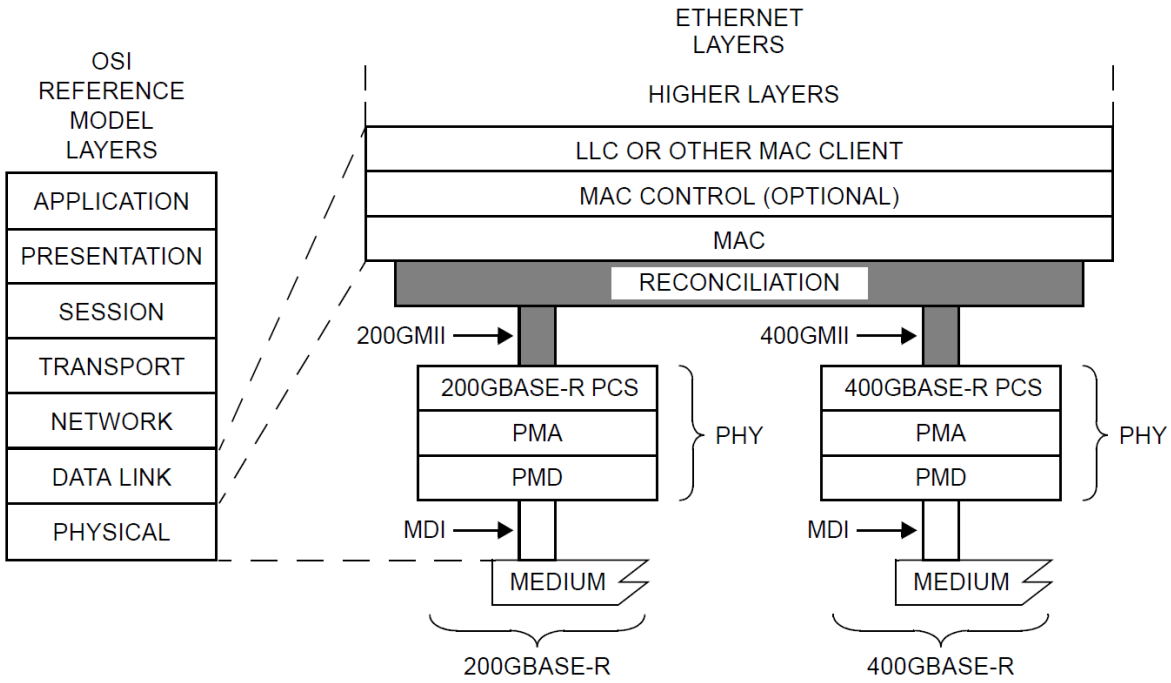
[a] The symbol ≈ indicates approximate equivalent of maximum Skew Variation in UI based on 1 UI equals 37.64706ps at PMD lane signaling rate of 26.5625GBd.

[b] The symbol ≈ indicates approximate equivalent of maximum Skew Variation in UI based on 1 UI equals 18.82353ps at PMD lane signaling rate of 53.125GBd.

Table 3-6. *Summary of Skew Variation constraints.*

3.2 Reconciliation Sublayer (RS) and Media Independent Interface (MII)

The RS and the MII are between Ethernet media access controllers and various PHYs. **Figure 3-7** shows the relationship of the RS and MII to the ISO/IEC OSI reference model.



200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
LLC = LOGICAL LINK CONTROL
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 3-7. *RS and MII in IEEE 802.3 Ethernet model.*

MII provides a logical interconnection between the MAC sublayer and PHY entities. MII is not intended to be physically instantiated, rather it can logically connect layers within a device.

The 200GMII supports 200Gb/s operation and the 400GMII supports 400Gb/s operation through its 64-bit-wide transmit and receive data paths. The RS provides a mapping between the signals provided at the MII (200GMII and 400GMII) and the MAC/PLS service definition.

The 200GMII and the 400GMII are optional logical interfaces between the MAC sublayer and the PHY. The 200GXS/400GXS sublayer in conjunction with the 200GAUI-n/400GAUI-n interface may be used to optionally extend the 200GMII/400GMII.

The RS adapts the bit serial protocols of the MAC to the parallel format of the PCS service interface. The PCS is specified to the 200GMII/400GMII, so if not implemented, a conforming implementation shall behave functionally as if the RS and 200GMII/400GMII were implemented.

The 200GMII/400GMII have the following characteristics:

a. The 200GMII supports a speed of 200Gb/s.
b. The 400GMII supports a speed of 400Gb/s.
c. Data and delimiters are synchronous to a clock reference.
d. They provide independent 64-bit wide transmit and receive data paths.
e. They support full duplex operation only.

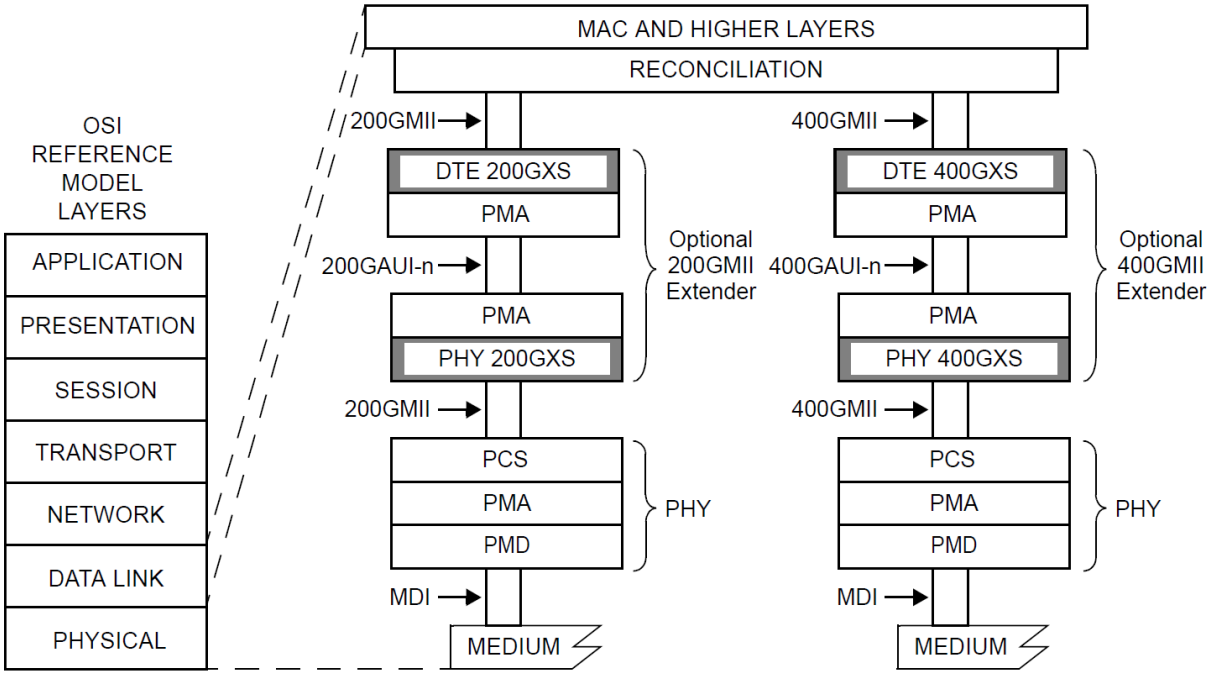
The following are the major concepts of the 200GMII/400GMII:

a. The 200GMII/400GMII are functionally similar to other MIIs that have been defined for lower speeds, as they all define an interface allowing independent development of MAC and PHY logic.
b. The RS converts between the MAC serial data stream and the parallel data paths of the 200GMII/400GMII.
c. The RS maps the signal set provided at the 200GMII/400GMII to the PLS service primitives provided at the MAC.
d. Each direction of data transfer is independent and serviced by data, control, and clock signals.
e. The RS generates continuous data or control characters on the transmit path and expects continuous data or control characters on the receive path.
f. The RS participates in link fault detection and reporting by monitoring the receive path for status reports that indicate an unreliable link and generating status reports on the transmit path to report detected link faults to the DTE on the remote end of the connecting link.
g. The 200GMII/400GMII may also support LPI signaling for PHY types supporting EEE.
h. The 200GMII/400GMII can be extended through the use of a pair of 200GXS/400GXS sublayers (DTE XS and PHY XS) with a 200GAUI-n/400GAUI-n between them.

3.3 200GMII Extender, 400GMII Extender, 200GMII Extender Sublayer (200GXS), and 400GMII Extender Sublayer (400GXS)

The 200GMII/400GMII Extender allows the extension of the 200GMII/400GMII to the PCS via a physical instantiation. The 200GMII/400GMII Extender is composed of a 200GXS/400GXS at the RS end, a 200GXS/400GXS at the PHY end with a physical instantiation of 200GAUI-n/400GAUI-n between two adjacent PMA sublayers. **Figure 3-8** shows the relationship of the 200GMII/400GMII Extender and 200GXS/400GXS sublayer with other sublayers to the ISO/IEC OSI reference model.

A 200GMII/400GMII Extender with the optional EEE capability encodes and decodes LPI signals. The assertion of LPI at the 200GMII/400GMII is encoded in the transmitted symbols. Detection of LPI encoding in the received symbols is indicated as LPI at the 200GMII/400GMII.



200GAUI-n = 200 Gb/s n-LANE ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
200GXS = 200GMII EXTENDER SUBLAYER
400GAUI-n = 400 Gb/s n-LANE ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
400GXS = 400GMII EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 3-8. *200GXS and 400GXS in IEEE 802.3 Ethernet model.*

The following is a list of the major concepts of the 200GMII/400GMII Extender:

a. Simple signal mapping to the 200GMII/400GMII.
b. The optional 200GMII/400GMII Extender can be inserted between the RS sublayer and the PHY to transparently extend the reach of the 200GMII/400GMII.
c. Independent transmit and receive data paths.
d. The 200GXS/400GXS leverages all functions and supports physical instantiations of the 200GAUI-n/400GAUI-n.
e. Optionally extends LPI signaling to the PHY for EEE.

The 200GXS, if implemented, shall be identical in function to the 200GBASE-R PCS with the addition of the functions based on a specific FEC degrade signaling. A single device may be configured as either a 200GXS or the 200GBASE-R PCS and may be managed through different optional management registers.

The 400GXS, if implemented, shall be identical in function to the 400GBASE-R PCS with the addition of the functions based on a specific FEC degrade signaling. A single device may be configured as either a 400GXS or the 400GBASE-R PCS and may be managed through different optional management registers.

A 200GMII Extender may use any of the following physical instantiations of the 200GAUI-n:

- 200GAUI-8 chip-to-chip.
- 200GAUI-8 chip-to-module.
- 200GAUI-4 chip-to-chip.
- 200GAUI-4 chip-to-module.

A 400GMII Extender may use any of the following physical instantiations of the 400GAUI-n:

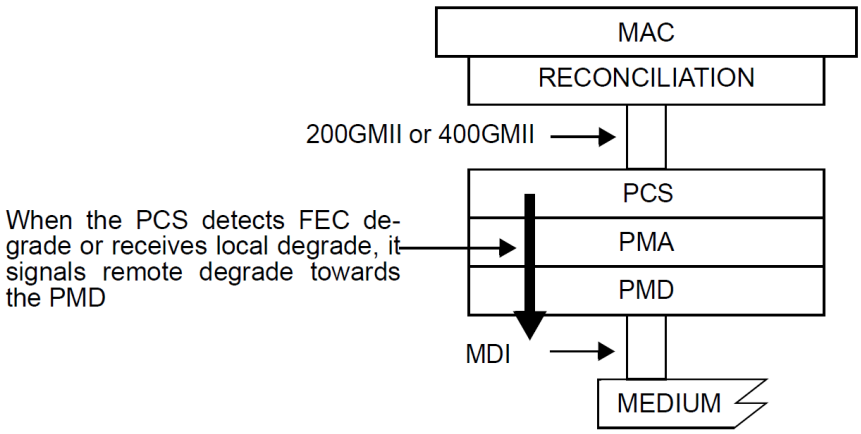
- 400GAUI-16 chip-to-chip.
- 400GAUI-16 chip-to-module.
- 400GAUI-8 chip-to-chip.
- 400GAUI-8 chip-to-module.

3.3.1 FEC Degrade

FEC degrade is an optional feature allowing for the detection of a non-service affecting link degradation condition based on exceeding a threshold for FEC corrected errors. If there are multiple FEC decoders in a given direction of transmission between the MAC sublayers at each end of the link, a Local Degrade condition is cascaded in

that direction of transmission to convey the fact that one or more FEC decoders in the path have exceeded their threshold of FEC corrected errors. If any FEC decoder in a given direction of transmission exceeds its provisioned threshold for FEC corrected errors, a Remote Degrade condition is indicated in the opposite direction of transmission from the PCS or eXtender Sublayer (XS) closest to the MAC.

Figure 3-9 illustrates the signaling of the Remote Degraded condition in the case that there is no XS present between the MAC and the PCS. Note that the PCS will not initiate the signaling for local degrade in this configuration as there are no additional FEC decoders in the receive direction between the PCS and the MAC.



200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 3-9. *Remote Degrade signaling without XS.*

Figure 3-10 illustrates the signaling of the Local Degraded condition in the case that an XS is present between the MAC and the PCS.

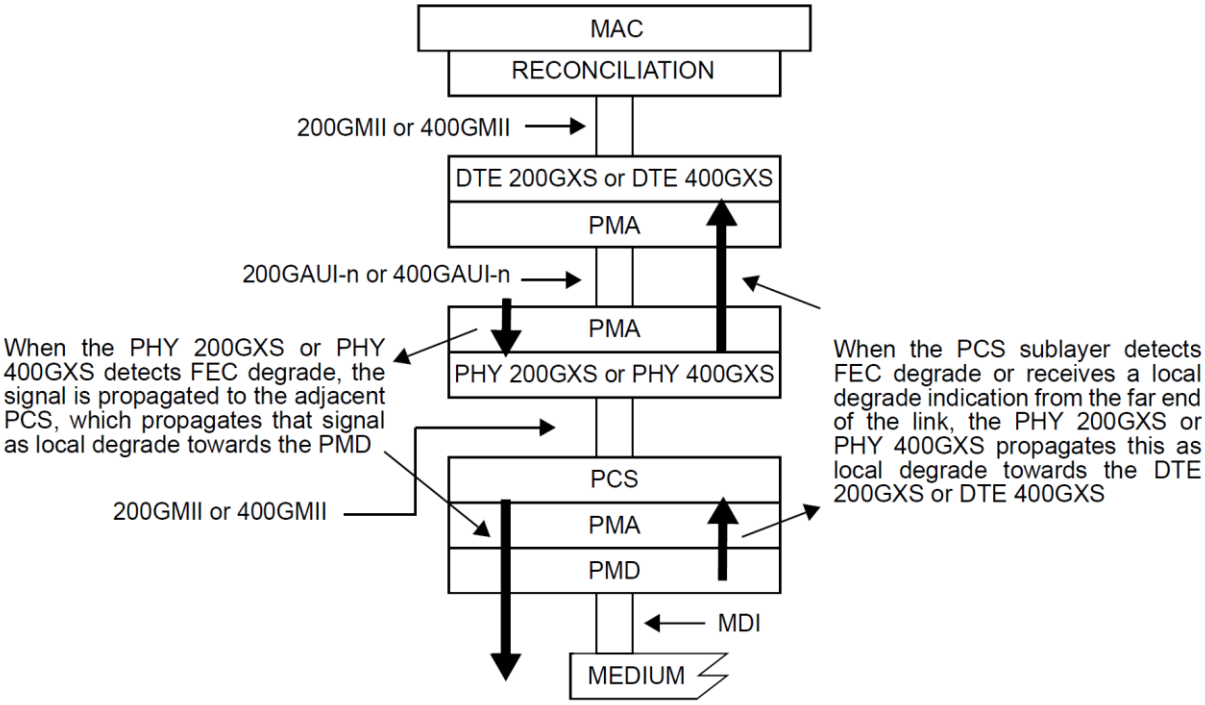
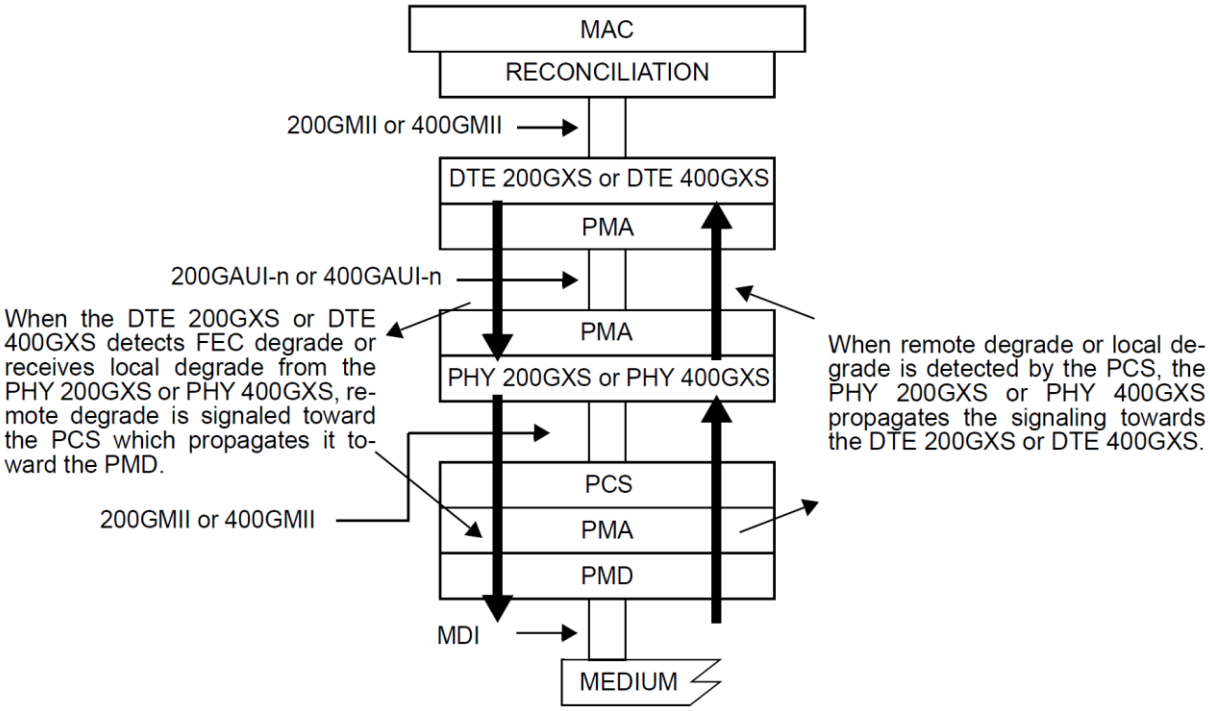


Figure 3-10. *Local Degrade signaling with XS.*

Figure 3-11 illustrates the signaling of the Remote Degraded condition in the case that an XS is present between the MAC and the PCS.



200GAUI-n = 200 Gb/s n-LANE ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
200GXS = 200GMII EXTENDER SUBLAYER
400GAUI-n = 400 Gb/s n-LANE ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
400GXS = 400GMII EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 3-11. *Remote Degrade signaling with XS.*

3.4 Physical Coding Sublayer (PCS) for 64B/66B, type 200GBASE-R and 400GBASE-R

The 200GBASE-R Physical Coding Sublayer (PCS) is a sublayer of the 200 Gb/s PHYs listed in **Table 3-1**. The 400GBASE-R PCS is a sublayer of the 400 Gb/s PHYs listed in **Table 3-2**. The terms 200GBASE-R and 400GBASE-R are used when referring generally to PHYs using the PCS. Both 200GBASE-R and 400GBASE-R are based on a 64B/66B code, which supports the transmission of data and control characters. The 64B/66B code is then transcoded to 256B/257B encoding to reduce the overhead and make room for FEC. The 256B/257B encoded data is then FEC encoded before being transmitted. Data distribution is introduced to support multiple lanes in the PHY, part of which includes the periodic insertion of an alignment marker, allowing the receive PCS to align data from multiple lanes.

Figure 3-12 depicts the relationship between the 200GBASE-R and 400GBASE-R sublayers (shown shaded), the Ethernet MAC and RS, and the higher layers.

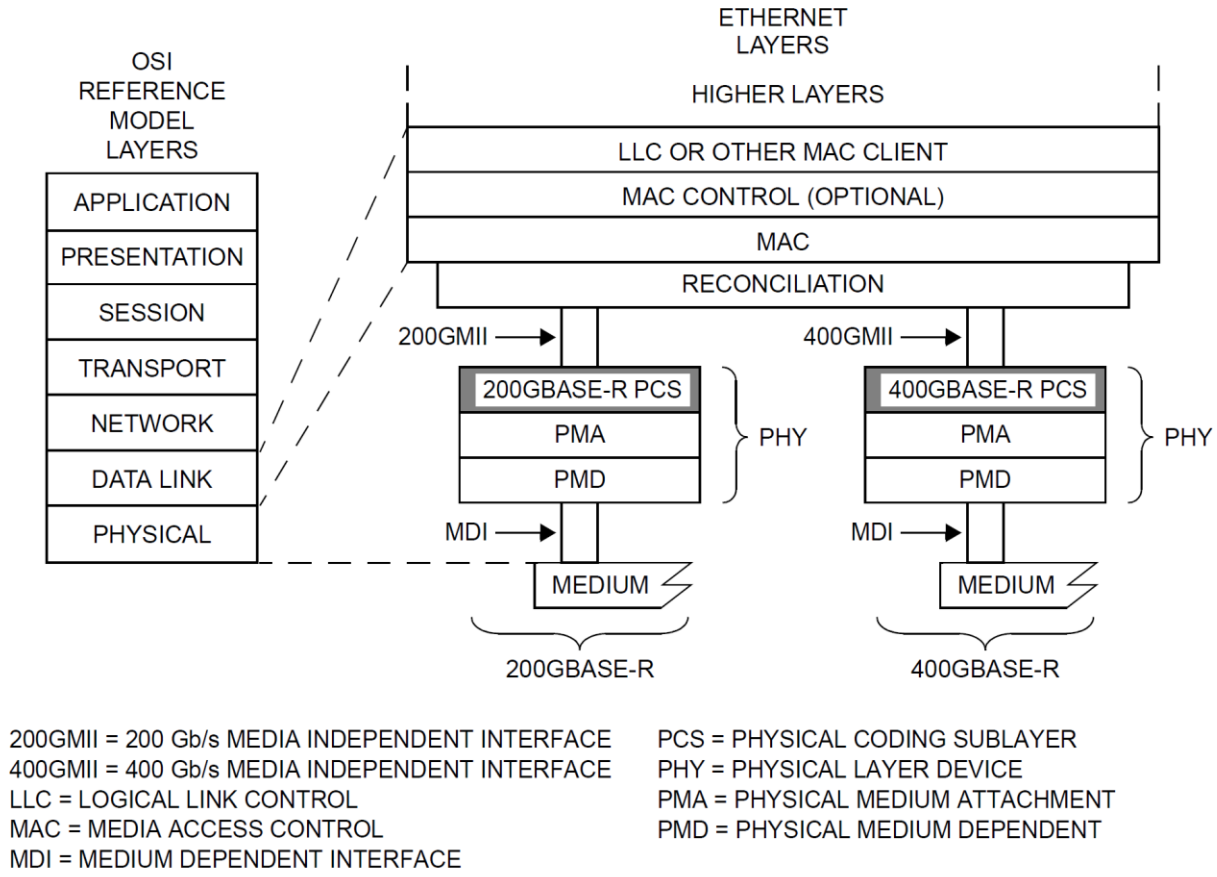


Figure 3-12. *200GBASE-R and 400GBASE-R PCS.*

The PCS service interface is the MII (200GMII/400GMII). The 200GMII provides a uniform interface to the RS for all 200 Gb/s PHY implementations. The 400GMII provides a uniform interface to the RS for all 400 Gb/s PHY implementations.

The 200GBASE-R and 400GBASE-R PCSs provide all services required by the 200GMII/400GMII, including the following:

a. Encoding (decoding) of eight 200GMII/400GMII data octets to (from) 66-bit blocks (64B/66B).
b. Transcoding from 66-bit blocks to (from) 257-bit blocks.
c. Reed-Solomon encoding (decoding) the 257-bit blocks.
d. Transferring encoded data to (from) the PMA.
e. Compensation for any rate differences caused by the insertion or deletion of alignment markers or due to any rate difference between the 200GMII/400GMII and PMA through the insertion or deletion of idle control characters.

f. Determining when a functional link has been established and informing the management entity via the MDIO when the PHY is ready for use.

The upper interface of the PCS connects to the RS sublayer through the 200GMII/400GMII. The lower interface of the PCS connects to the PMA sublayer to support a PMD. The 200GBASE-R PCS has a nominal rate at the PMA service interface of 26.5625Gtransfers/s on each of 8 PCS lanes, which provides capacity for the MAC data rate of 200 Gb/s. The 400GBASE-R PCS has a nominal rate at the PMA service interface of 26.5625Gtransfers/s on each of 16 PCS lanes, which provides capacity for the MAC data rate of 400 Gb/s.

The PCS service interface allows the 200GBASE-R or 400GBASE-R PCS to transfer information to and from a PCS client, which is the RS sublayer.

The PMA Service Interface supports the exchange of the encoded data between the PCS and the PMA sublayer.

3.5 Physical Medium Attachment (PMA) sublayer, type 200GBASE-R and 400GBASE-R

The PMA sublayer allows the PCS to connect in a media-independent way with a range of physical media, or for the DTE XS to connect to the PHY XS. The 200GBASE-R PMA(s) can support any of the 200 Gb/s PMDs in **Table 3-1**, and the 400GBASE-R PMA(s) can support any of the 400 Gb/s PMDs in **Table 3-2**. The terms 200GBASE-R and 400GBASE-R are used when referring generally to PHYs using the PMA. 200GBASE-R and 400GBASE-R can be extended to support any full duplex medium requiring only that the PMD be compliant with the appropriate PMA interface.

For 200GBASE-R PMAs, electrical interfaces connecting PMA sublayers, known as 200GAUI-n, are defined in **Annex B**, **Annex C**, **Annex D**, and **Annex E**. For 400GBASE-R PMAs, electrical interfaces connecting PMA sublayers, known as 400GAUI-n, are defined in **Annex B**, **Annex C**, **Annex D**, and **Annex E**.

Figure 3-13 shows the PMA sublayer (shown shaded) with the other sublayers to the ISO/IEC OSI reference model.

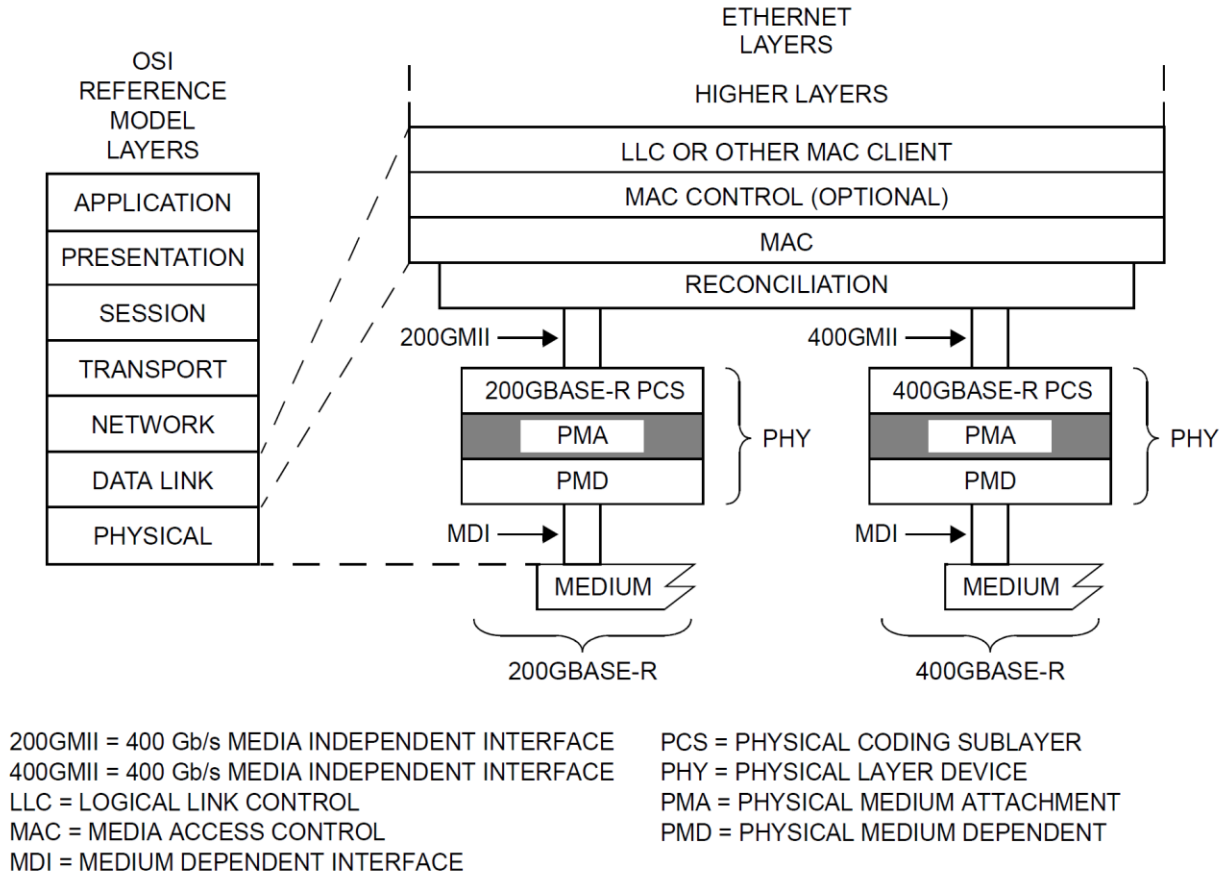


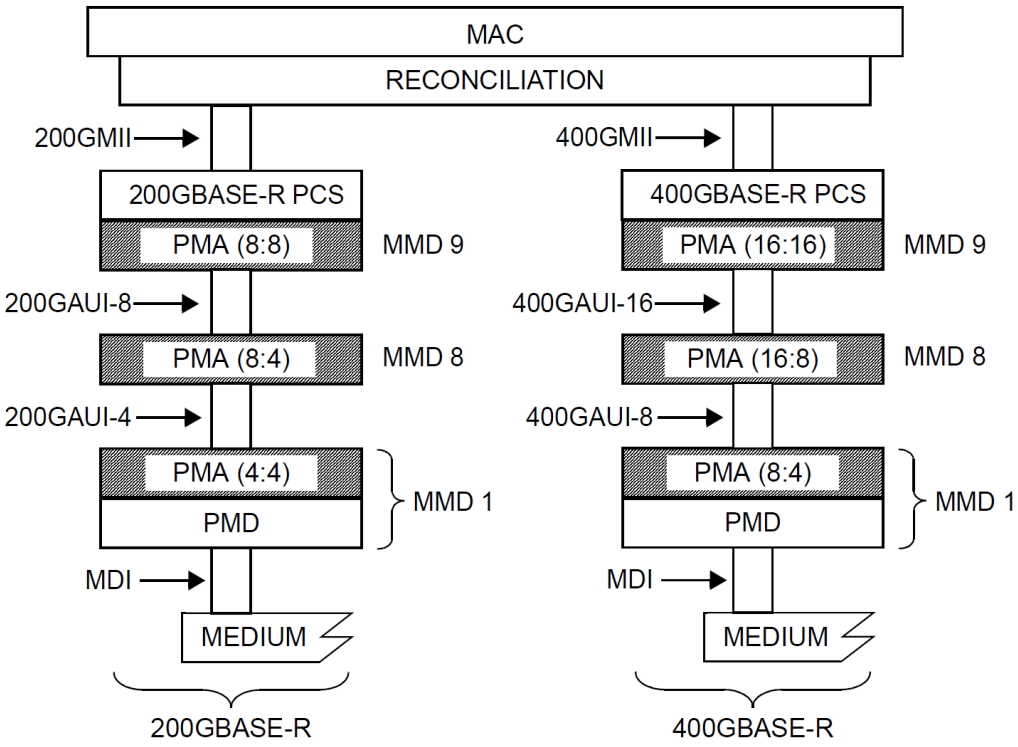
Figure 3-13. *200GBASE-R and 400GBASE-R PMA.*

The following is a summary of the principal functions implemented (when required) by the PMA in both the transmit and receive directions:

a. Adapt the Physical Coding Sublayer Lane (PCSL) formatted signal to the appropriate number of abstract or physical lanes.
b. Provide per input-lane clock and data recovery.
c. Provide bit-level multiplexing.
d. Provide clock generation.
e. Provide signal drivers.
f. Optionally provide local loopback to/from the PMA service interface.
g. Optionally provide remote loopback to/from the PMD service interface.
h. Optionally provide test-pattern generation and detection.
i. Tolerate Skew Variation.
j. Perform PAM4 encoding and decoding for 200GBASE-R PMAs where the number of physical lanes is 4, and for 400GBASE-R PMAs where the number of physical lanes is 4 or 8.

In addition, the PMA provides receive link status information in the receive direction.

An implementation may use one or more PMA sublayers to adapt the number and rate of the PCS lanes to the number and rate of the PMD lanes. The number of PMA sublayers required depends on the partitioning of functionality for a particular implementation. An example is illustrated in **Figure 3-14**. Additional examples are illustrated in **Annex A**. Each PMA maps the PCSLs from *p* PMA input lanes to *q* PMA output lanes in the *Tx* direction, and from *q* PMA input lanes to *p* PMA output lanes in the *Rx* direction.



200GAUI = 200 Gb/s ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI = 400 Gb/s ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 3-14. *Example 200GBASE-R and 400GBASE-R PMA layering.*

MDIO Manageable Device (MMD) addresses 1, 8, 9, and 10 are available for addressing multiple instances of PMA sublayers (see **Table 3-7** for MMD device address). If the PMA sublayer that is closest to the PMD is packaged with the PMD, it shares MMD 1 with the PMD. More addressable instances of PMA sublayers, each one separated from lower addressable instances by chip-to-chip interfaces, may be implemented and addressed allocating MMD addresses to PMAs in increasing

numerical order going from the PMD toward the MAC. The example shown in **Figure 3-14** could be implemented with three addressable instances: MMD 1 addressing the lowest PMA sublayer (co-packaged with the PMD), MMD 8 addressing the PMA sublayer above the 200GAUI-4 below the 200GAUI-8 or above the 400GAUI-8 below the 400GAUI-16, and MMD 9 addressing the PMA sublayer below the PCS.

The number of input lanes and the number of output lanes for a PMA are always divisors of the number of PCSLs. For PMA sublayers supporting 200GBASE-R PMDs, the number of PCSLs is 8. For PMA sublayers supporting 400GBASE-R PMDs, the number of PCSLs is 16.

Device address	MMD name
0	Reserved
1	PMA/PMD
2	WIS
3	PCS
4	PHY XS
5	DTE XS
6	TC
7	Auto-Negotiation
8	Separated PMA (1)
9	Separated PMA (2)
10	Separated PMA (3)
11	Separated PMA (4)
12	OFDM PMA/PMD
13	Power Unit
14 through 28	Reserved
29	RS and MII extension
30	Vendor specific 1
31	Vendor specific 2

Table 3-7. *MDIO Manageable Device addresses.*

The following guidelines apply to the partitioning of PMAs:

a. The inter-sublayer service interface is used for the PMA service interfaces supporting a flexible architecture with multiple PMA sublayers.
b. An instance of this interface can only connect service interfaces with the same number of lanes, where the lanes operate at the same rate.

c. 200GAUI-n is a physical instantiation of the connection between two adjacent 200GBASE-R PMA sublayers with the exception of the *inst:IS_SIGNAL.indication* which is carried outside of this physically instantiated interface. 400GAUI-n is a physical instantiation of the connection between two adjacent 400GBASE-R PMA sublayers with the exception of the *inst:IS_SIGNAL.indication* which is carried outside of this physically instantiated interface.
    i. As physical instantiations, these define electrical and timing specification as well as requiring a receive re-timing function.
    ii. 200GAUI-8 is a 26.5625GBd by 8 lane NRZ physical instantiation of the 200Gb/s connection. 400GAUI-16 is a 26.5625GBd by 16 lane NRZ physical instantiation of the 400Gb/s connection.
    iii. 200GAUI-4 is a 26.5625GBd by 4 lane PAM4 physical instantiation of the 200Gb/s connection. 400GAUI-8 is a 26.5625GBd by 8 lane PAM4 physical instantiation of the 400Gb/s connection.
d. The abstract inter-sublayer service interface can be physically instantiated as a 200GAUI-n or 400GAUI-n, using associated PMAs to map to the appropriate number of lanes.
e. Opportunities for optional test-pattern generation, optional test-pattern detection, optional local loopback and optional remote loopback are dependent upon the location of the PMA sublayer in the implementation.
f. A minimum of one PMA sublayer is required in a PHY.
g. A maximum of four PMA sublayers are addressable as MDIO MMDs.

3.6 Physical Medium Dependent (PMD) sublayer and medium, type 200GBASE-DR4

200GBASE-DR4 PMD is specified together with the single-mode fiber medium. The optical signal generated by this PMD type is modulated using a 4-level pulse amplitude modulation (PAM4) format. When forming a complete PHY, a PMD shall be connected to the appropriate PMA as shown in **Table 3-8**, to the medium through the MDI and optionally with the management functions that may be accessible through the management interface (MDIO), or equivalent.

Associated Section	200GBASE-DR4
3.2 – RS	Required
3.2 – 200GMII[a]	Optional
3.3 – 200GMII Extender	Optional
3.4 – PCS	Required
3.5 – PMA	Required
Annex B – Chip-to-chip 200GAUI-8	Optional
Annex C – Chip-to-module 200GAUI-8	Optional
Annex D – Chip-to-chip 200GAUI-4	Optional
Annex E – Chip-to-chip 200GAUI-4	Optional
Annex F – Energy Efficient Ethernet	Optional
[a] The 200GMII is an optional interface. However, if the 200GMII is not implemented, a conforming implementation must behave functionally as though the RS and 200GMII were present.	

Table 3-8. *PHY paragraphs associated with the 200GBASE-DR4 PMD.*

Figure 3-15 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC OSI reference model. 200 Gb/s Ethernet is introduced in Section **3**.

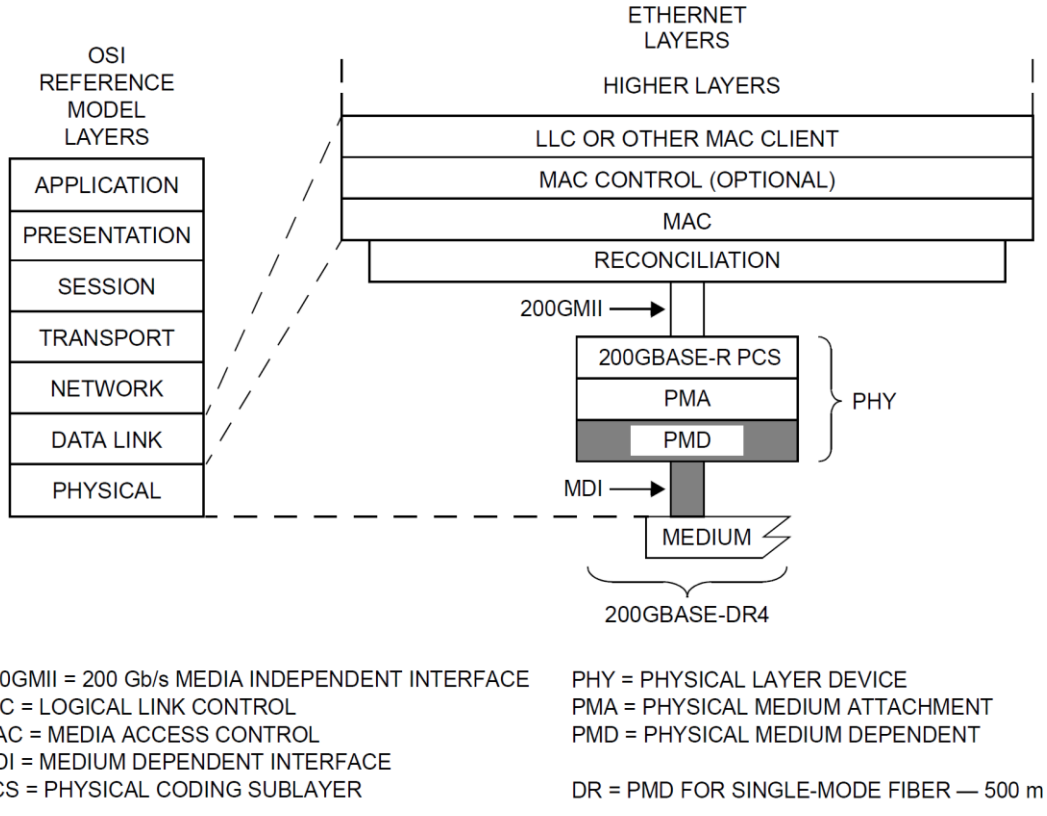


Figure 3-15. *200GBASE-DR4 PMD.*

200GBASE-DR4 PHYs with the optional EEE fast wake capability may enter the LPI mode to conserve energy during periods of low link utilization (see **Annex F**). The deep sleep mode of EEE is not supported.

The BER when processed according to Section **3.5** shall be less than 2.4×10^{-4} provided that the error statistics are sufficiently random that this results in a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap when processed according to Section **3.5** and then Section **3.4**. For a complete PHY, the frame loss ratio may be degraded to 6.2×10^{-11} for 64-octet frames with minimum interpacket gap due to additional errors from the electrical interfaces.

If the error statistics are not sufficiently random to meet this requirement, then the BER shall be less than that required to give a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap.

The sum of the transmit and receive delays at one end of the link contributed by the 200GBASE-DR4 PMD including 2m of fiber in one direction shall be no more than 4096-bit times (8 *pause_quanta* or 20.48 ns). A description of overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **3.1.3** and its references.

The Skew (relative delay between the lanes) and Skew Variation must be kept within limits so that the information on the lanes can be reassembled by the PCS. Skew and Skew Variation are defined in Section **3.1.3** and specified at the points SP1 to SP6 shown in **Figure 3-5** and **Figure 3-6**.

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

3.6.1 PMD Functional Specifications

The 200GBASE-DR4 PMD performs the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

The PMD block diagram is shown in **Figure 3-16**. The optical transmit signal is defined at the output end of a single-mode fiber patch cord (TP2), between 2m and 5m in length. Unless specified otherwise, all transmitter measurements and tests defined in Section **3.6.4** are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see Section **3.6.7.1**). Unless specified otherwise, all receiver measurements and tests defined in Section **3.6.4** are made at TP3.

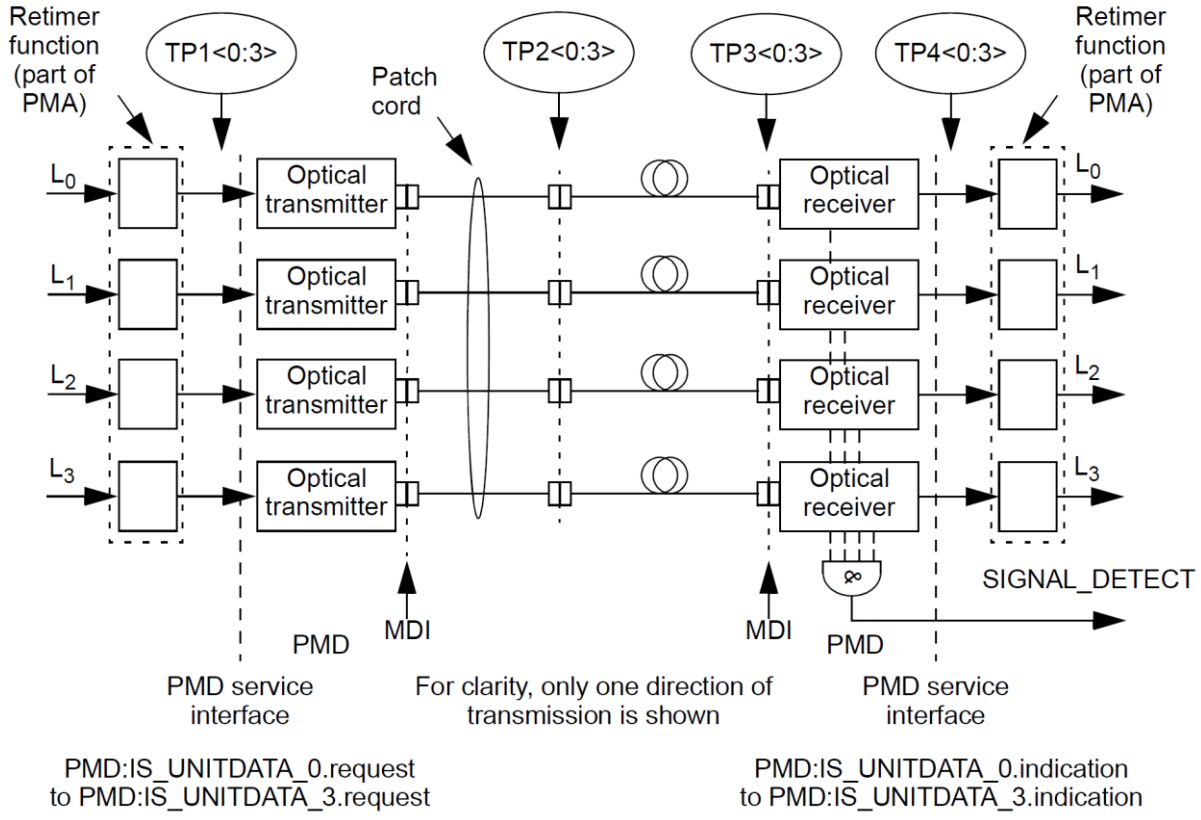


Figure 3-16. *Block diagram for 200GBASE-DR4 transmit/receive paths.*

3.6.1.1 PMD transmit function

The PMD Transmit function shall convert the four symbol streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_3.request* into four separate optical signals. The four optical signals shall then be delivered to the MDI, which contains four parallel light paths for transmit, according to the transmit optical specifications in this section. The highest

optical power level in each signal stream shall correspond to *tx_symbol = three* and the lowest shall correspond to *tx_symbol = zero*.

3.6.1.2 PMD receive function

The PMD Receive function shall convert the four parallel optical signals received from the MDI into separate symbol streams for delivery to the PMD service interface using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_3.indication*, all according to the receive optical specifications in this section. The higher optical power level in each signal stream shall correspond to *rx_symbol = three* and the lowest shall correspond to *rx_symbol = zero*.

3.6.1.3 PMD global signal detect function

The PMD global signal detect function reports the state of the SIGNAL_DETECT via the PMD service interface. The *SIGNAL_DETECT* parameter is signaled continuously, while the *PMD:IS_SIGNAL.indication* message is generated when a change occurs in the value of *SIGNAL_DETECT*.

SIGNAL_DETECT is a global indicator of the presence of optical signals on all four lanes. The value of the *SIGNAL_DETECT* parameter shall be generated according to the conditions defined in **Table 3-9**. The PMD receiver is not required to verify whether a compliant 200GBASE-DR4 signal is being received.

Receive conditions	***SIGNAL_DETECT* value**
For any lane. Average optical power at TP3 ≤ –16dBm	FAIL
For all lanes. *[(Optical power at TP3 ≥ average receive power, each lane (min) **Table 3-12**)* AND (Compliant 200GBASE–R signal input)]	OK
All other conditions	Unspecified

Table 3-9. *SIGNAL_DETECT value definition.*

All implementations must provide adequate margin between the input optical power level at which the *SIGNAL_DETECT* parameter is set to OK, and the inherent noise level of the PMD, including the effects of the crosstalk, power supply noise, etc.

Various implementations of the Signal Detect function are permitted by this standard (IEEE Std 802.3bs), including implementations that generate the *SIGNAL_DETECT*

parameter values in response to the amplitude of the modulation of the optical signal and implementations that respond to the average optical power of the modulated optical signal.

3.6.1.4 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard (IEEE Std 802.3bs). When the MDIO is implemented, each *PMD_signal_detect_i*, where *i* represents the lane number in the range 0:3, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of **Table 3-9**.

3.6.2 Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for 200GBASE-DR4. While it is expected that a PMD will map electrical lane *i* to optical lane *i* and vice versa, there is no need to define the physical ordering of the lanes, as the PCS sublayer is capable of receiving the lanes in any arrangement. The positioning of transmit and receive lanes at the MDI is specified in Section **3.6.7.2**.

3.6.3 PMD to MDI optical specifications for 200GBASE-DR4

The operating range for the 200GBASE-DR4 PMD is defined in **Table 3-10**. A 200GBASE-DR4 compliant PMD operates on type B1.1, B1.3, or B6_a single-mode fibers according to the specifications defined in **Table 3-18**. A PMD that exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., a 200GBASE-DR4 PMD operating at 600m meets the operating range requirement of 2 m to 500 m).

PMD type	Required operating range
200GBASE-DR4	2m to 500m

Table 3-10. *200GBASE-DR4 operating range.*

3.6.3.1 200GBASE-DR4 transmitter optical specifications

The 200GBASE-DR4 transmitter shall meet the specifications defined in **Table 3-11** per the definitions in Section **3.6.4**.

Description	Value	Unit
Signaling rate, each lane (range)	26.5625 ± 100 ppm	GBd
Modulation format	PAM4	-
Lane wavelength (range)	1304.5 to 1317.5	nm

Side-mode suppression ratio (SMSR), (min)	30	dB
Average launch power, each lane (max)	3	dBm
Average launch power, each lane [a] (min)	-5.1	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (max)	2.8	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (min) [b]	-3	dBm
Launch power in OMA_{outer} minus TDECQ, each lane (min)	-4.4	dBm
Transmitter and dispersion eye closure for PAM4 (TDECQ), each lane (max)	3.4	dB
Average launch power of OFF transmitter, each lane (max)	-16	dBm
Extinction ratio, each lane (min)	3.5	dB
$RIN_{21.4}OMA$ (max)	-132	dB/Hz
Optical return loss tolerance (max)	21.4	dB
Transmitter reflectance [c] (max)	-26	dB
[a] Average launch power, each lane (min) is informative and not the principal indicator of signal strength. [b] Even if the TDECQ < 1.4 dB, the OMA_{outer} (min) must exceed this value. [c] Transmitter reflectance is defined looking into the transmitter.		

Table 3-11. *200GBASE-DR4 transmit characteristics.*

3.6.3.2 200GBASE-DR4 receive optical specifications

The 200GBASE-DR4 receiver shall meet the specifications defined in **Table 3-12** per the definitions in Section **3.6.4**.

Description	Value	Unit
Signaling rate, each lane (range)	26.5625 ± 100 ppm	GBd
Modulation format	PAM4	-
Lane wavelengths (range)	1304.5 to 1317.5	Nm
Damage threshold [a], each lane	4	dBm
Average receive power, each lane (max)	3	dBm
Average receive power, each lane [b] (min)	-8.1	dBm
Receive power (OMA_{outer}), each lane (max)	2.8	dBm
Receiver reflectance (max)	-26	dB
Receiver sensitivity (OMA_{outer}), each lane [c] (max)	-6.6	dBm
Stressed receiver sensitivity (OMA_{outer}), each lane [d] (max)	-4.1	dBm

Conditions of stressed receiver sensitivity test: [e]		
- Stressed eye closure for PAM4 (SECQ), lane under test	3.4	dB
- OMA_{outer} of each aggressor lane	2.8	dBm
[a] The receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level. [b] Average receive power, each lane (min) is informative and not the principal indicator of signal strength. [c] Receiver sensitivity (OMA_{outer}), each lane (max) is informative and is defined for a transmitter with SECQ of 0.9dB. [d] Measured with conformance test signal at TP3 for the BER specified in Section **3.6**. [e] These test conditions are for measuring stressed receiver sensitivity. They are not characteristics of the receiver.		

Table 3-12. *200GBASE-DR4 receive characteristics.*

3.6.3.3 200GBASE-DR4 illustrative link power budget

An illustrative power budget and penalties for 200GBASE-DR4 channels are shown in **Table 3-13**.

Parameter	Value	Unit
Power budget (for max TDECQ)	6.5	dB
Operating distance	500	m
Channel insertion loss [a]	3	dB
Maximum discrete reflectance	See Section **3.6.7**	dB
Allocation for penalties [b] (for max TDECQ)	3.5	dB
Additional insertion loss allowed	0	dB
[a] The channel insertion loss is calculated using the maximum distance specified in **Table 3-10** and cabled optical fiber attenuation of 0.5dB/km at 1304.5nm plus an allocation for connection and splice loss given in Section **3.6.7**. [b] Link penalties are used for link budget calculations.		

Table 3-13. *200GBASE-DR4 illustrative link power budget.*

3.6.4 Definition of optical parameters and measurements methods

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified.

3.6.4.1 Test patterns for optical parameters

While compliance must be achieved in normal operation, specific test patterns are defined for measurement consistency and the measurement of some parameters.

Table 3-15 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subsections in which each parameter is defined. Any of the test patterns given for a particular test in **Table 3-15** can be used to perform that test. The test patterns used in this Section are shown in **Table 3-14**.

Pattern	Pattern description
Square wave	Square wave (8 threes, 8 zeros)
3	PRBS31Q
4	PRBS13Q
5	Scrambled idle
6	SSPRQ

Table 3-14. *Test patterns.*

Parameter	Pattern	Related Sections
Wavelength	Square wave, 3, 4, 5, 6 or valid 200GBASE-R signal	**3.6.4.2**
Side mode suppression ratio	3, 5, 6 or valid 200GBASE-R signal	**-**
Average optical power	3, 5, 6 or valid 200GBASE-R signal	**3.6.4.3**
Outer Optical Modulation Amplitude (OMA_{outer})	4 or 6	**3.6.4.4**
Transmitter and dispersion eye closure for PAM4 (TDECQ)	6	**3.6.4.5**
Extinction ratio	4 or 6	**3.6.4.10**
RIN21.4OMA	Square wave	**3.6.4.11**
Stressed receiver sensitivity	3 or 5	**3.6.4.13**

Table 3-15. *Test-pattern definitions and related sections.*

3.6.4.2 Wavelength and side mode suppression ratio (SMSR)

The wavelength of each optical lane shall be within the range given in **Table 3-10** if measured per IEC 61280-1-3. The lane under test is modulated using the test pattern defined in **Table 3-15**.

3.6.4.3 Average optical power

The average optical power of each lane shall be within the limits given in **Table 3-10** if measured using the methods given in IEC 61280-1-1. The average optical power is

measured using the test pattern defined in **Table 3-15**, per the test setup in **Figure 3-17**.

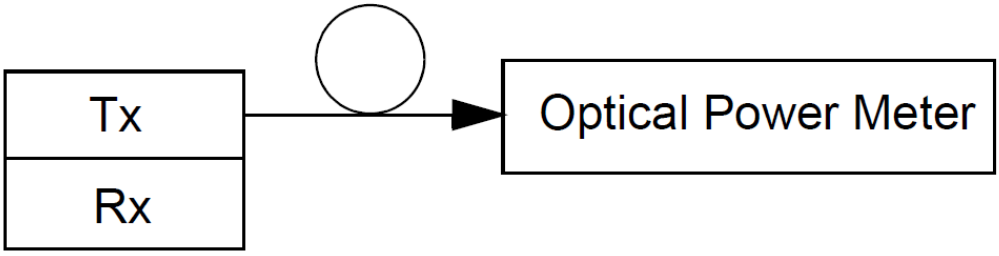


Figure 3-17. *Optical power measurement test set-up.*

3.6.4.4 Outer Optical Modulation Amplitude (OMA_{outer})

The OMA_{outer} of each lane shall be within the limits given in **Table 3-10**. The OMA_{outer} is measured using a test pattern specified for OMA_{outer} in **Table 3-15** as the difference between the average optical launch power level P_3, measured over the central 2 UI of a run of 7 threes, and the average optical launch power level P_0, measured over the central 2 UI of a run of 6 zeros, as shown in **Figure 3-18**.

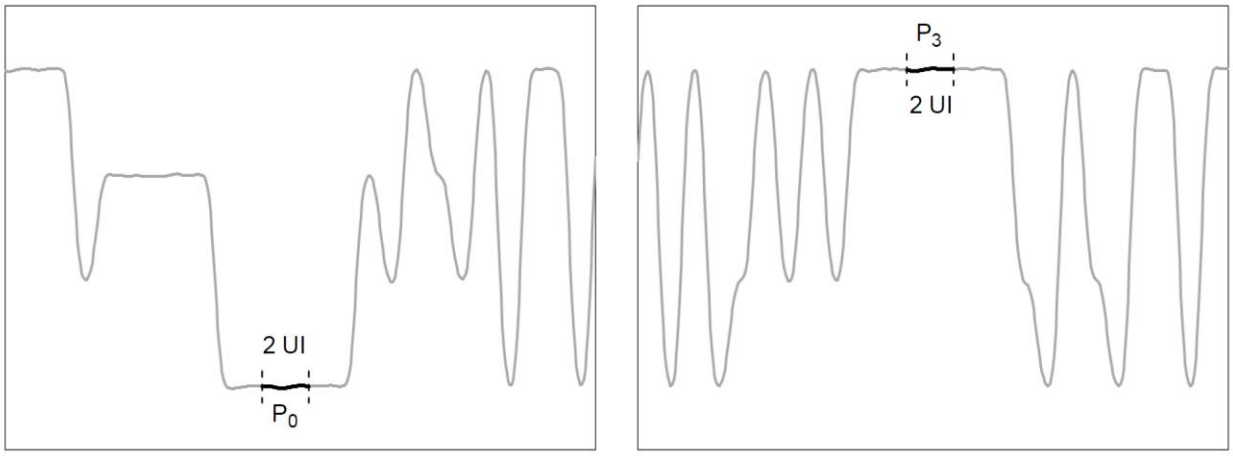


Figure 3-18. *Example power levels P_0 and P_3 from PRBS13Q test pattern.*

3.6.4.5 Transmitter and dispersion eye closure for PAM4 (TDECQ)

The TDECQ of each lane shall be within the limits given in **Table 3-10** if measured using the methods specified in Sections **3.6.4.6**, **3.6.4.7**, and **3.6.4.8**.

TDECQ is a measure of each optical transmitter's vertical eye closure when transmitted through a worst-case optical channel (specified in Section **3.6.4.7**), as measured through an optical to electrical converter (O/E) and oscilloscope with the combined frequency response given in Section **3.6.4.8**, and equalized with the

reference equalizer (as described in Section **3.6.4.9**). The reference receiver and equalizer may be implemented in software or may be part of an oscilloscope.

Table 3-15 specifies the test patterns to be used for measurement of TDECQ.

3.6.4.6 TDECQ conformance test setup

A block diagram for the TDECQ conformance test is shown in **Figure 3-19**. Each optical lane is tested individually with all other lanes in operation and all lanes using the same test pattern. There shall be at least 31 UI delay between the test pattern on one lane and the pattern on any other lane, so that the symbols on each lane are not correlated within the PMD. The optical splitter and variable reflector are adjusted so that each transmitter is tested with the optical return loss specified in **Table 3-16**. The state of polarization of the back reflection is adjusted to create the greatest RIN. Each optical lane is tested with the optical channel described in Section **3.6.4.7**. The combination of the O/E and the oscilloscope has a fourth-order Bessel-Thomson filter response with a bandwidth of approximately 13.28125GHz. Compensation can be made for any deviation from an ideal fourth-order Bessel-Thomson response.

The test pattern (specified in **Table 3-15**) is transmitted repetitively by the optical lane under test and the oscilloscope is set up to capture the complete pattern for TDECQ analysis as described in Section **3.6.4.8**. The clock recovery unit (CRU) has a corner frequency of 4MHz and a slope of 20dB/decade. The CRU can be implemented in hardware or software depending on oscilloscope technology.

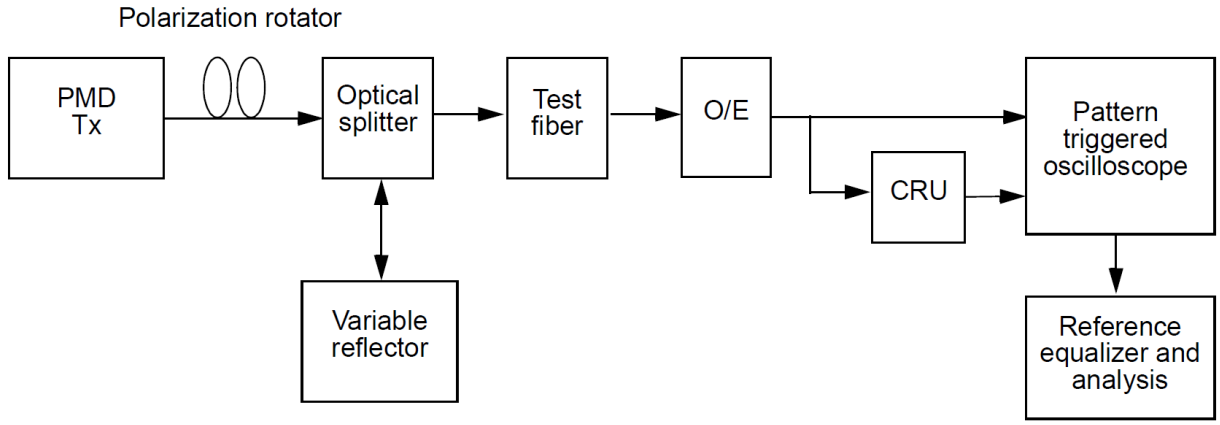


Figure 3-19. *TDECQ conformance test block diagram.*

3.6.4.7 Channel requirements

The transmitter is tested using an optical channel that meets the requirements listed in **Table 3-16**.

PMD type	Dispersion [a] (ps/nm)		Insertion loss [b]	Optical return loss [c]	Max mean DGD
	Minimum	Maximum			
200GBASE-DR4	$0.011625\times\lambda\times[1-(1324/\lambda)^4]$	$0.011625\times\lambda\times[1-(1300/\lambda)^4]$	Minimum	21.4dB	0.5ps

[a] The dispersion is measured for the wavelength of the device under test (λ in nm). The coefficient assumes 500m for 200GBASE-DR4.

[b] There is no intent to stress the sensitivity of the O/E converter associated with the oscilloscope.

[c] The optical return loss is applied at TP2.

Table 3-16. *Transmitter compliance channel specifications.*

A 200GBASE-DR4 transmitter shall be compliant with a total dispersion of at least as negative as the “minimum dispersion” and at least as positive as the “maximum dispersion” columns specified in **Table 3-16** for the wavelength of the device under test. This may be achieved with channels consisting of fibers with lengths chosen to meet the dispersion requirements.

To verify the correct amount of dispersion of the fiber, the measurement method defined in IEC 60793-1-42 must be used. The measurement is made in the linear power regime of the fiber.

The channel provides an optical return loss specified in **Table 3-16**. The polarization of the back reflection is adjusted to create the greatest Relative Intensity Noise (RIN).

The mean Differential Group Delay (DGD) of the channel is to be less than the value specified in **Table 3-16**.

3.6.4.8 TDECQ measurement method

The standard deviation of the noise of the O/E and oscilloscope combination, σ_S, is defined with no optical input signal and the same settings as used to capture the histograms described below.

OMA_{outer} is measured according to Section **3.6.4.4** on the equalized signal. The test pattern specified for TDECQ (see **Table 3-15**) is transmitted repetitively by the optical lane under test and the oscilloscope is set up to capture samples from all symbols in the complete pattern without averaging.

If an equivalent-time sampling oscilloscope is used, the impact of the sampling process and the reference equalizer on transmitter noise must be compensated for, so that the correct magnitude of noise is present at the output of the equalizer.

The captured waveform is processed to reach the largest noise that could be combined with the signal of an ideal reference receiver, when it optimally equalized by a reference equalizer. The optimal equalizer tap coefficients depend on the amount of noise that can be added to the signal, so the process of finding the noise that can be added and the optimal equalizer setting is iterative. One way to do this, is to use an estimated PAM4 symbol error ratio, as the figure of merit for the equalized signal described below.

The reference equalizer (specified in Section **3.6.4.9**) is applied to the waveform. The sum of the equalizer tap coefficients is equal to 1. The eye diagram is formed from the equalized captured waveform.

The average optical power (P_{ave}) of the equalized eye diagram is defined, and the 0 UI and 1 UI crossing points are defined by the average of the eye diagram crossing times, as measured at P_{ave}, as illustrated in **Figure 3-20**.

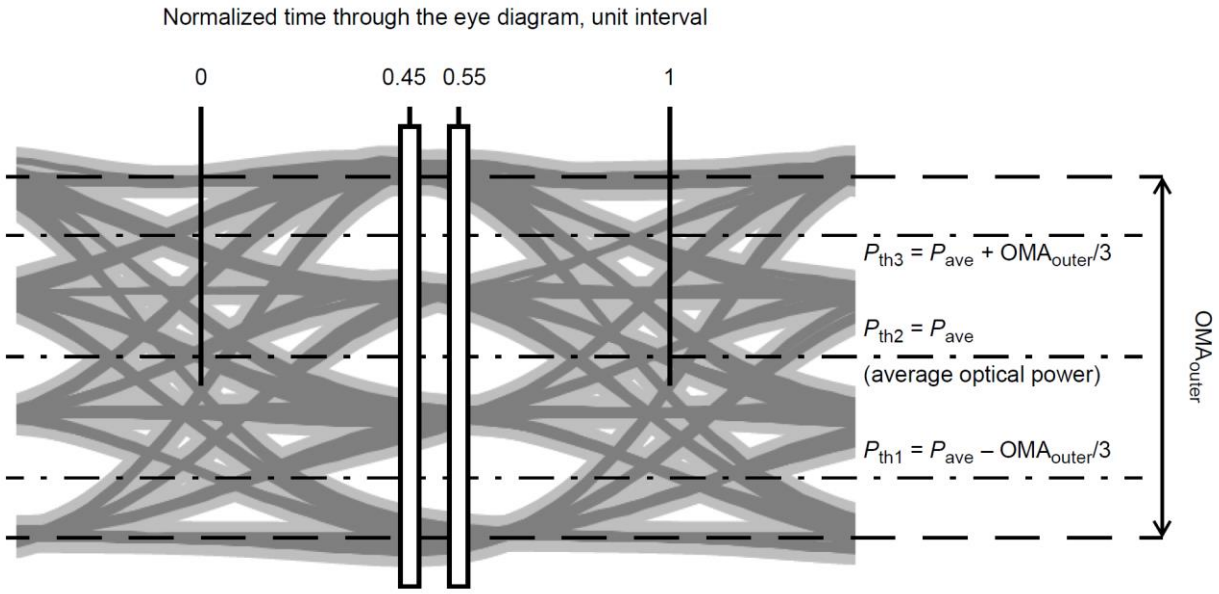


Figure 3-20. *Illustration of the TDECQ measurement.*

Two vertical histograms are measured through the eye diagram, nominally centered at 0.45 UI and 0.55 UI. Each of the histogram windows spans all of the modulation levels of the eye diagram, as illustrated in **Figure 3-20**. The precise time position of the pair of histograms is adjusted to minimize TDECQ while keeping the histograms spaced 0.1 UI apart.

Each histogram window has a width of 0.04 UI. Each histogram window has outer height boundaries which are set beyond the extremes of the eye diagram.

3.6.4.9 TDECQ reference equalizer

The reference equalizer for 200GBASE-DR4 is a 5 tap, T spaced, Feed-Forward Equalizer (FFE), where T is the symbol period. A functional model of the reference equalizer is shown in **Figure 3-21**. The sum of the equalizer tap coefficients is equal to 1. The tap 1, tap 2, or tap 3, has the largest magnitude tap coefficient.

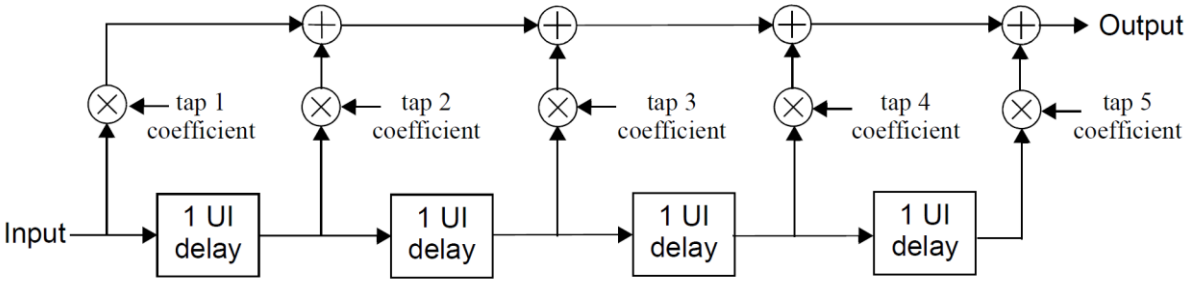


Figure 3-21. *TDECQ reference equalizer functional model*

NOTE: This reference equalizer is part of the TDECQ test and does not imply any particular receiver equalizer implementation.

3.6.4.10 Extinction ratio

The extinction ratio of each lane shall be within the limits given in **Table 3-11** if measured using a test pattern specified for extinction ratio in **Table 3-15**. The extinction ratio of a PAM4 optical signal is defined as the ratio of the average optical launch power level P_3, measured over the central 2 UI of a run of 7 threes, and the average optical launch power level P_0, measured over the central 2 UI of a run of 6 zeros, as shown in **Figure 3-18**.

3.6.4.11 Relative intensity noise ($RIN_{21.4}OMA$)

RIN shall be as defined by the measurement methodology of Relative Intensity Noise Optical Modulation Amplitude (RIN_xOMA) measuring procedure with the following exceptions:

a. The optical return loss is 21.4dB.
b. Each lane may be tested individually with the sum of the optical power from all of the lanes not under test being below -30dBm.
c. The upper -3dB limit of the measurement apparatus is to be approximately equal to the signaling rate (i.e., 26.6GHz).
d. The test pattern is according to **Table 3-15**.

3.6.4.12 Receiver sensitivity

Receiver sensitivity, which is defined for an input signal with SECQ of 0.9dB (e.g., an ideal input signal without overshoot), is informative and compliance is not required. If measured, the test signal should have negligible impairments, such as the InterSymbol Interference (ISI), the rise/fall times, the jitter, and the RIN. Instead, the normative requirement for receivers is stressed receiver sensitivity.

3.6.4.13. Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in **Table 3-12**, using the test pattern specified for stressed receiver sensitivity in **Table 3-15**. The BER is required to be met for the lane under test on its own.

Stressed receiver sensitivity is defined with all transmit and receive lanes in operation. Any of the patterns specified for stressed receiver sensitivity in **Table 3-15** is sent from the transmit section of the PMD under test. The signal being transmitted is asynchronous to the received signal.

3.6.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

200GBASE-DR4 optical transceivers shall conform to Hazard Level 1 laser requirements as defined in IEC 60825-1 and IEC 60825-2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service.

It is recommended that proper installation practices, as defined by applicable local codes and regulation, be followed.

Normative specifications in this section shall be met by a system integrating a 200GBASE-DR4 PMD over the life of the product while the product operates within the manufacturer's range of environmental, power, and other specifications.

It is recommended that manufacturers indicate in the literature associated with the PHY the operating environmental conditions to facilitate selection, installation, and maintenance. It is recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of this section will be met.

A system integrating a 200GBASE-DR4 PMD shall comply with applicable local and national codes for the limitation of electromagnetic interference.

3.6.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 3-22**.

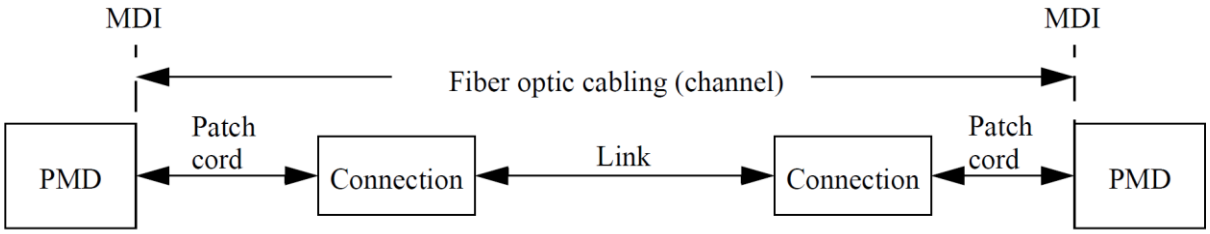


Figure 3-22. *Fiber optic cabling model.*

The channel insertion loss is given in **Table 3-17**. A channel contains additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion comply with the specifications. Insertion loss measurements of installed fiber cables are made in accordance with IEC 6128-1-1. The fiber optic cabling model (channel) determined here is the same as a simplex fiber optic permanent link. The term channel is used here for consistency with generic cabling standards.

Description	200GBASE-DR4	Unit
Operating distance (max)	500	m
Channel insertion loss [a], [b] (max)	3	dB
Channel insertion loss (min)	0	dB
Positive dispersion [b] (max)	0.8	ps/nm
Negative dispersion [b] (min)	-0.93	ps/nm
DGD_max [c]	2.24	ps
Optical return loss (min)	37	dB
[a] These channel insertion loss values include cable, connectors, and splices. [b] Over the wavelength range 1304.5nm to 1317.5nm. [c] Differential Group Delay (DGD) is the time difference at reception between the fractions of a pulse that were transmitted in the two principal states of polarization of an optical signal. *DGD_max* is the maximum differential group delay that the system must tolerate.		

Table 3-17. *Fiber optic cabling (channel) characteristics.*

3.6.7 Characteristics of the fiber optic cabling (channel)

The 200GBASE-DR4 fiber optic cabling shall meet the specifications defined in **Table 3-18**. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

The fiber optic cable requirements are satisfied by cables containing IEC 60793-2-50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-mode), or type B6_a (bend insensitive) fibers or the requirements in **Table 3-18** where they differ.

Description	Value	Unit
Nominal fiber specification wavelength	1310	nm
Cabled optical fiber attenuation (max)	0.4 [a]	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0.093	ps/nm^2km
[a] The 0.4dB/km attenuation is provided for Outside Plant cable as defined in CENELEC EN 50173-1 and ISO/IEC 11801-1.		

Table 3-18. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 3-22**, consists of a mated pair of optical connectors.

The maximum link distance is based on an allocation of 2.75dB total connection and splice loss. For example, this allocation supports five connections with an average insertion loss per connection of 0.4dB. Connections with different loss characteristics may be used provided the requirements of **Table 3-17** are met.

The maximum value for each discrete reflectance shall be less than or equal to the value shown in **Table 3-19** corresponding to the number of discrete reflectances above –55dB within the channel. For numbers of discrete reflectances in between the two numbers shown in the table, the lower of the two corresponding maximum discrete reflectance values applies.

Number of discrete reflectances above –55 dB	Maximum value for each discrete reflectance
1	-37dB
2	-42dB
4	-45dB
6	-47dB

8	-48dB
10	-49dB

Table 3-19. *Maximum value of each discrete reflectance.*

3.6.7.1 Medium Dependent Interface (MDI)

The 200GBASE-DR4 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the "fiber optic cabling" (as shown in **Figure 3-22**). The 200GBASE-DR4 PMD is coupled to the fiber optic cabling through one connector plug into the MDI optical connector as shown in **Figure 3-24**. Example constructions of the MDI include the following:

a. PMD with a connectorized fiber pigtail plugged into an adapter.
b. PMD connector.

3.6.7.2 Optical lane assignments

The four transmit and four receive optical lanes of 200GBASE-DR4 shall occupy the positions depicted in **Figure 3-23** when looking into the MDI connector with the keyway feature on top. The interface contains eight active lanes within twelve total positions. The transmit optical lanes occupy the left-most four positions and the receive optical lanes occupy the right-most four positions. The four center positions are unused.

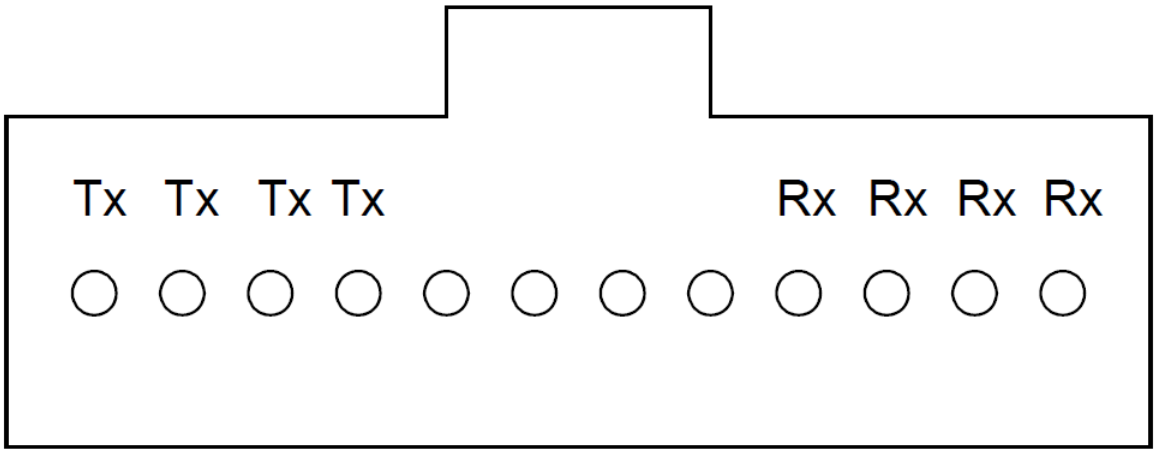


Figure 3-23. *200GBASE-DR4 optical lane assignments.*

3.6.7.3 MDI requirements

The MDI shall meet the dimensional specifications of IEC 61754-7-1 interface 7-1-9: MPO device connector, angled interface. The plug terminating the optical fiber cabling shall meet the dimensional specifications of IEC 61754-7-1 interface 7-1-1: MPO female plug connector, down-angled interface for 2 to 12 fibres. The MDI shall

optically mate with the plug on the optical fiber cabling. **Figure 3-24** shows an MPO female plug connector with down-angled interface, and an MDI as an active device connector with angled interface.

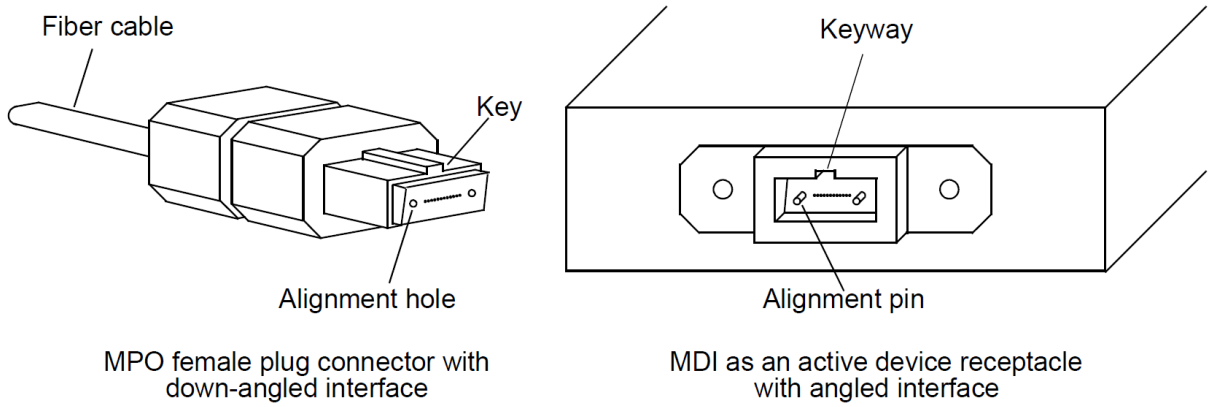


Figure 3-24. *MPO female plug with down-angled interface and MDI active device connector with angled interface.*

The MDI shall meet the interface performance specifications of IEC 61753-021-2 for performance level D/2.

NOTE: Transmitter compliance testing is performed at TP2 as defined in Section **3.6.1**, not at the MDI.

3.7 Physical Medium Dependent (PMD) sublayer and medium, type 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, 400GBASE-LR8

200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and the 400GBASE-LR8 PMDs are specified together with the single-mode fiber medium. The optical signals generated by these four PMD types are modulated using a 4-level pulse amplitude modulation (PAM4) format. When forming a complete PHY, a PMD shall be connected to the appropriate PMA as shown in **Table 3-20**, to the medium through the MDI and optionally with the management functions that may be accessible through the management interface (MDIO), or equivalent.

Associated section	200GBASE-FR4, 200GBASE-LR4	400GBASE-FR8, 400GBASE-LR8
3.2 – RS	Required	Required
3.2 – 200GMII [a]	Optional	Not applicable
3.2 – 400GMII [a]	Not applicable	Optional
3.3 – 200GMII Extender	Optional	Not applicable
3.3 – 400GMII Extender	Not applicable	Optional

3.4 – PCS for 200GBASE-R	Required	Not applicable
3.4 – PCS for 400GBASE-R	Not applicable	Required
3.5 – PMA for 200GBASE-R	Required	Not applicable
3.5 – PMA for 400GBASE-R	Not applicable	Required
Annex B – Chip-to-chip 200GAUI-8	Optional	Not applicable
Annex B – Chip-to-chip 400GAUI-16	Not applicable	Optional
Annex C – Chip-to-module 200GAUI-8	Optional	Not applicable
Annex C – Chip-to-module 400GAUI-16	Not applicable	Optional
Annex D – Chip-to-chip 200GAUI-4	Optional	Not applicable
Annex D – Chip-to-chip 400GAUI-8	Not applicable	Optional
Annex E – Chip-to-module 200GAUI-4	Optional	Not applicable
Annex E – Chip-to-module 400GAUI-8	Not applicable	Optional
Annex F – Energy Efficiency Ethernet	Optional	Optional
[a] 200GMII and 400GMII are optional interfaces. However, if the appropriate interface is not implemented, a conforming implementation must behave functionally as though the RS and 200GMII or 400GMII were present.		

Table 3-20. *Physical Layer sections associated with the 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 PMDs.*

Figure 3-25 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC OSI reference model. 200 Gb/s and 400 Gb/s Ethernet are introduced in Section **3**.

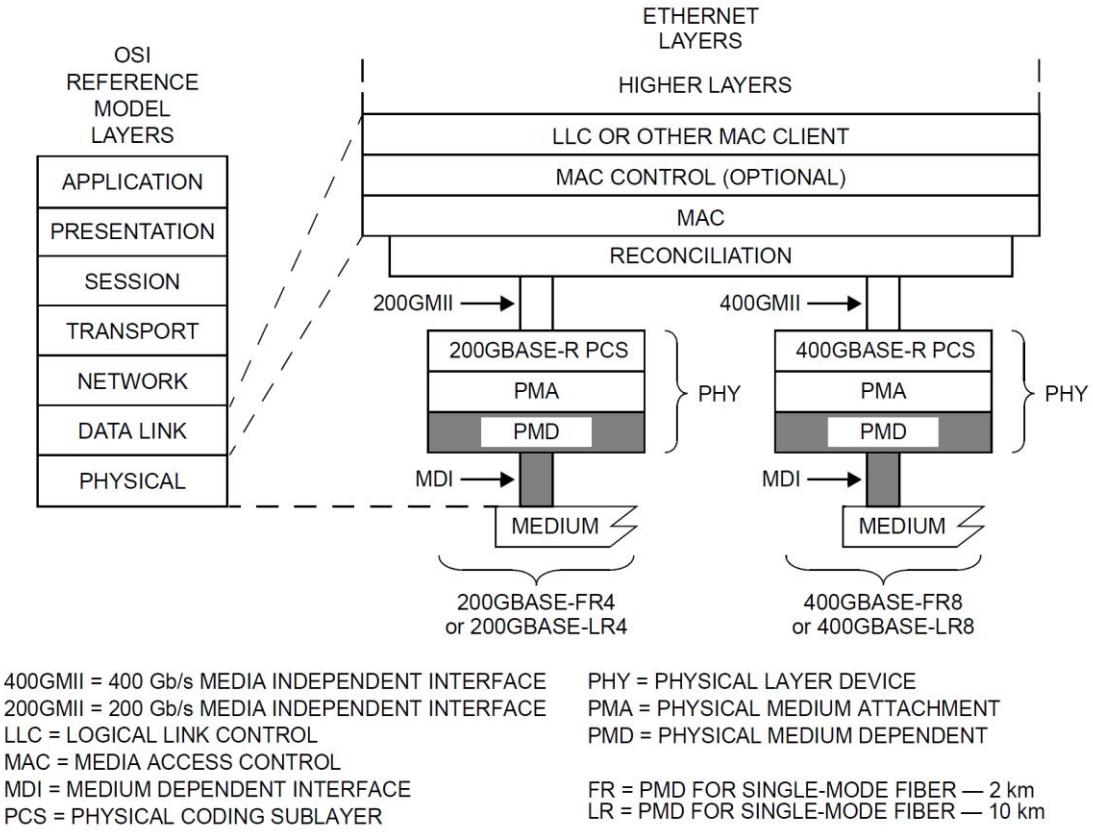


Figure 3-25. *200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 PMDs.*

200GBASE-FR4 and 200GBASE-LR4 use four lanes, while 400GBASE-FR8 and 400GBASE-LR8 use eight lanes. In this section, where there are four or eight items (depending on PMD type) such as lanes, the items are numbered from *0* to *n-1*, and an example item is numbered *i*. Thus *n* is 4 or 8.

200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 PHYs with the optional EEE fast wake capability may enter the LPI mode to conserve energy during periods of low link utilization (see **Annex F**). The deep sleep mode of EEE is not supported.

The BER when processed according to Section **3.5** shall be less than 2.4×10^{-4} provided that the error statistics are sufficiently random that this results in a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap when processed according to Section **3.5** and then Section **3.4**. For a complete PHY, the frame loss ratio may be degraded to 6.2×10^{-11} for 64-octet frames with minimum interpacket gap due to additional errors from the electrical interfaces.

If the error statistics are not sufficiently random to meet this requirement, then the BER shall be less than that required to give a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap.

The sum of the transmit and receive delays at one end of the link contributed by the 200GBASE-FR4 or 200GBASE-LR4 PMD including 2m of fiber in one direction shall be no more than 4096-bit times (8 *pause_quanta* or 20.48ns). The sum of the transmit and receive delays at one end of the link contributed by the 400GBASE-FR8 or 400GBASE-LR8 PMD including 2m of fiber in one direction shall be no more than 8192-bit times (16 *pause_quanta* or 20.48ns). A description of overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **3.1.3** and its references.

The Skew (relative delay between the lanes) and Skew Variation must be kept within limits so that the information on the lanes can be reassembled by the PCS. Skew and Skew Variation are defined in Section **3.1.4** and specified at the points SP1 to SP6 shown in **Figure 3-5** and **Figure 3-6**.

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

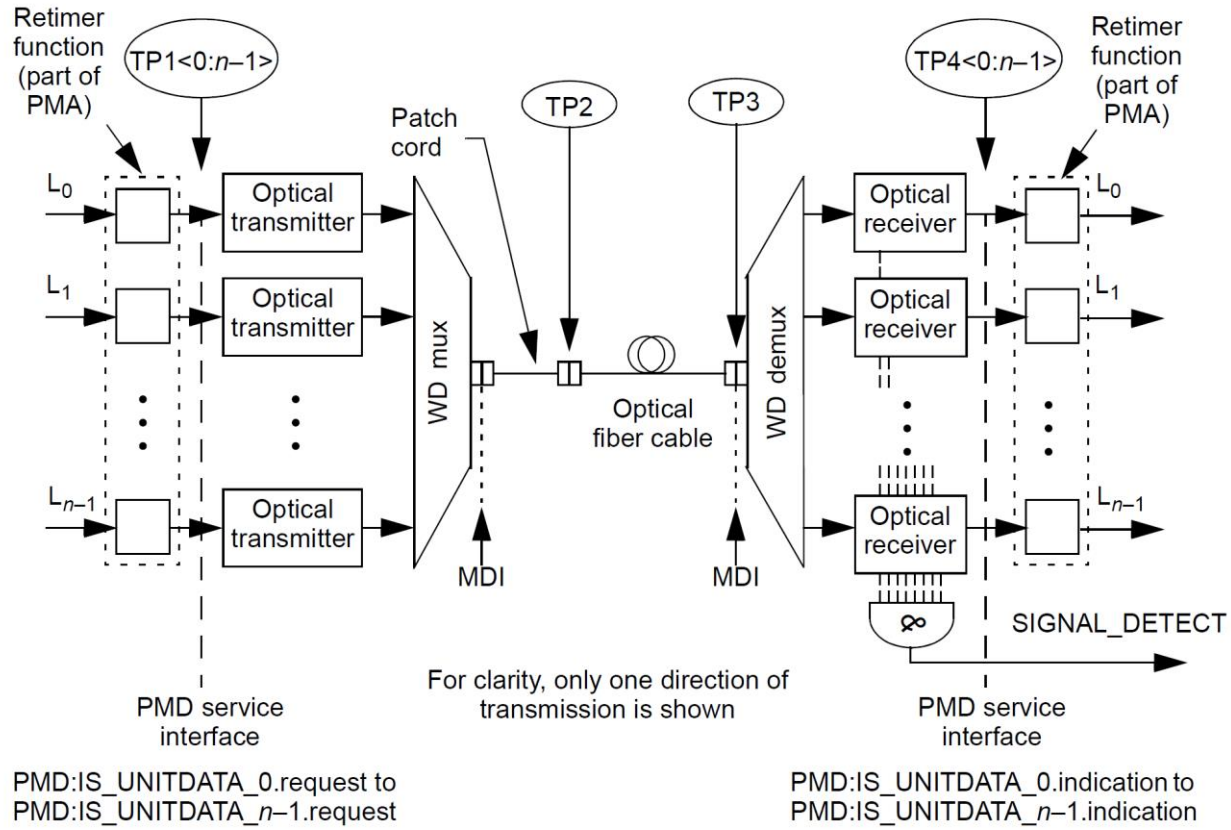
If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

3.7.1 PMD Functional Specifications

The 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 PMDs perform the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

The PMD block diagram is shown in **Figure 3-26**. The optical transmit signal is defined at the output end of a single-mode fiber patch cord (TP2), between 2m and 5m in length. Unless specified otherwise, all transmitter measurements and tests defined in Section **3.7.4** are made at TP2. The optical receive signal is defined at the

output of the fiber optic cabling (TP3) at the MDI (see Section **3.7.7.1**). Unless specified otherwise, all receiver measurements and tests defined in Section **3.7.4** are made at TP3.



NOTE—Specification of the retimer function and the electrical implementation of the PMD service interface is beyond the scope of this standard.

Figure 3-26. *Block diagram for 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 transmit/receive paths.*

TP1<0:n–1> and *TP4<0:n–1>* are informative reference points that may be useful to implementers for testing components (these test points will not typically be accessible in an implemented system).

3.7.1.1 PMD transmit function

The PMD Transmit function shall convert the *n* symbol streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_n–1.request* into n separate optical signals. The *n* optical signals shall then be wavelength division multiplexed and delivered to the MDI, all according to the transmit optical specifications in this section. The highest optical power level in

each signal stream shall correspond to *tx_symbol = three* and the lowest shall correspond to *tx_symbol = zero*.

3.7.1.2 PMD receive function

The PMD Receive function shall demultiplex the composite optical signal received from the MDI into *n* separate optical signals. The *n* optical signals shall then be converted into *n* symbol streams for delivery to the PMD service interface using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_n–1.indication*, all according to the receive optical specifications in this section. The higher optical power level in each signal shall correspond to *rx_symbol = three* and the lowest shall correspond to *rx_symbol = zero*.

3.7.1.3 PMD global signal detect function

The PMD global signal detect function shall report the state of *SIGNAL_DETECT* via the PMD service interface. The *SIGNAL_DETECT* parameter is signaled continuously, while the *PMD:IS_SIGNAL.indication* message is generated when a change in the value of *SIGNAL_DETECT* occurs.

SIGNAL_DETECT shall be a global indicator of the presence of optical signals on all *n* lanes. The value of the *SIGNAL_DETECT* parameter shall be generated according to the conditions defined in **Table 3-21**. The PMD receiver is not required to verify whether a compliant 200GBASE-R or 400GBASE-R signal is being received. This standard imposes no response time requirements on the generation of the *SIGNAL_DETECT* parameter.

Receive conditions	***SIGNAL_DETECT* value**
For any lane. Average optical power at TP3 ≤ -30dBm	FAIL
For all lanes. [(Optical power at TP3 ≥ average receive power, each lane (min) in **Table 3-28** for 200GBASE-FR4 and 200GBASE-LR4 or **Table 3-29** for 400GBASE-FR8 and 400GBASE-LR8) AND (Compliant 200GBASE-R or 400GBASE–R signal input)]	OK
All other conditions	Unspecified

Table 3-21. *SIGNAL_DETECT value definition.*

As an unavoidable consequence of the requirements for the setting of the *SIGNAL_DETECT* parameter, implementations must provide adequate margin

between the input optical power level at which the *SIGNAL_DETECT* parameter is set to *OK*, and the inherent noise level of the PMD including the effects of crosstalk, power supply noise, etc.

Various implementations of the Signal Detect function are permitted by this standard (IEEE Std 802.3bs), including implementations that generate the *SIGNAL_DETECT* parameter values in response to the amplitude of the modulation of the optical signal and implementations that respond to the average optical power of the modulated optical signal.

3.7.1.4 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard. When the MDIO is implemented, each *PMD_signal_detect_i*, where *i* represents the lane number in the range 0:*n–1*, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of **Table 3-21**.

3.7.2 Wavelength-division-multiplexing lane assignments

The wavelength range for each lane of the 200GBASE-FR4 PMD is defined in **Table 3-22**. The 200GBASE-FR4 center wavelengths are members of the CWDM wavelength grid defined in ITU-T G.694.2 and are spaced at 20nm.

The wavelength range for each lane of the 200GBASE-LR4 PMD is defined in **Table 3-23**. The wavelength range for each lane of the 400GBASE-FR8 and 400GBASE-LR8 PMDs is defined in **Table 3-24**. The 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 center frequencies are members of the frequency grid for 100GHz spacing and above defined in ITU-T G.694.1 and are spaced at 800GHz.

Lane	Center wavelength	Wavelength range
L_0	1271nm	1264.5 to 1277.5nm
L_1	1291nm	1284.5 to 1297.5nm
L_2	1311nm	1304.5 to 1317.5nm
L_3	1331nm	1324.5 to 1337.5nm

Table 3-22. *200GBASE-FR4 wavelength-division-multiplexed lane assignments.*

Lane	Center frequency	Center wavelength	Wavelength range
L_0	231.4THz	1295.56nm	1294.53 to 1296.59nm
L_1	230.6THz	1300.05nm	1299.02 to 1301.09nm

L_2	229.8THz	1304.58nm	1303.54 to 1305.63nm
L_3	229THz	1309.14nm	1308.09 to 1310.19nm

Table 3-23. *200GBASE-LR4 wavelength-division-multiplexed lane assignments.*

NOTE: There is no requirement to associate a particular electrical lane with a particular optical lane, as the PCS is capable of receiving lanes in any arrangement.

Lane	Center frequency	Center wavelength	Wavelength range
L_0	235.4THz	1273.54nm	1272.55 to 1274.54nm
L_1	234.6THz	1277.89nm	1276.89 to 1278.89nm
L_2	233.8THz	1282.26nm	1281.25 to 1283.27nm
L_3	233THz	1286.66nm	1285.65 to 1287.68nm
L_4	231.4THz	1295.56nm	1294.53 to 1296.59nm
L_5	230.6THz	1300.05nm	1299.02 to 1301.09nm
L_6	229.8THz	1304.58nm	1303.54 to 1305.63nm
L_7	229THz	1309.14nm	1308.09 to 1310.19nm

Table 3-24. *400GBASE-FR8 and 400GBASE-LR8 wavelength-division-multiplexed lane assignments.*

3.7.3 PMD to MDI optical specifications for 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8

The operating ranges for the 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 PMDs are defined in **Table 3-25**. A 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, or 400GBASE-LR8 compliant PMD operates on type B1.1, B1.3, or B6_a single-mode fibers according to the specifications defined in **Table 3-35**. A PMD that exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., a 400GBASE-FR8 PMD operating at 2.5km meets the operating range requirement of 2m to 2km). The 400GBASE-LR8 PMD interoperates with the 400GBASE-FR8 PMD provided that the channel requirements for 400GBASE-FR8 are met.

PMD type	Required operating range
200GBASE-FR4 and 400GBASE-FR8	2m to 2km
200GBASE-LR4 and 400GBASE-LR8	2m to 10km

Table 3-25. *200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 operating ranges.*

3.7.3.1 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 transmitter optical specifications

The 200GBASE-FR4 transmitter shall meet the specifications defined in **Table 3-26** per the definitions in Section **3.7.4**. The 200GBASE-LR4 transmitter shall meet the specifications defined in **Table 3-26** per the definitions in Section **3.7.4**. The 400GBASE-FR8 transmitter shall meet the specifications defined in **Table 3-27** per the definitions in Section **3.7.4**. The 400GBASE-LR8 transmitter shall meet the specifications defined in **Table 3-27** per the definitions in Section **3.7.4**.

Description	200GBASE-FR4	200GBASE-LR4	Unit
Signaling rate, each lane (range)	26.5625 ± 100 ppm		GBd
Modulation format	PAM4		-
Lane wavelengths (range)	1264.5 to 1277.5 1284.5 to 1297.5 1304.5 to 1317.5 1324.5 to 1337.5	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19	nm
Side-mode suppression ratio (SMSR), (min)	30		dB
Total average launch power (max)	10.7	11.3	dBm
Average launch power, each lane (max)	4.7	5.3	dBm
Average launch power, each lane [a] (min)	-4.2	-3.4	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (max)	4.5	5.1	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (min) [b]	-1.2	-0.4	dBm
Difference in launch power between any two lanes (OMA_{outer}) (max)	4		dB
Launch power in OMA_{outer} minus TDECQ, each lane (min): for extinction ratio ≥ 4.5dB for extinction ratio < 4.5dB	 -2.6 -2.5	 -1.8 -1.7	 dBm dBm
Transmitter and dispersion eye closure for PAM4 (TDECQ), each lane (max)	3.3	3.4	dB
Average launch power of OFF transmitter, each lane (max)	-30		dBm
Extinction ratio, each lane (min)	3.5		dB

$RIN_{16.5}OMA$ (max)	-132	-	dB/Hz
$RIN_{15.1}OMA$ (max)	-	-132	dB/Hz
Optical return loss tolerance (max)	16.5	15.1	dB
Transmitter reflectance [c] (max)	-26		dB
[a] Average launch power, each lane (min) is informative and not the principal indicator of signal strength. A transmitter with launch power below this value cannot be compliant; however, a value above this does not ensure compliance. [b] Even if the TDECQ < 1.4dB for an extinction ratio of ≥ 4.5dB or TDECQ < 1.3dB for an extinction ratio of < 4.5dB, the OMA_{outer} (min) must exceed this value. [c] Transmitter reflectance is defined looking into the transmitter.			

Table 3-26. *200GBASE-FR4 and 200GBASE-LR4 transmit characteristics.*

Description	200GBASE-FR4	200GBASE-LR4	Unit
Signaling rate, each lane (range)	26.5625 ± 100 ppm		GBd
Modulation format	PAM4		-
Lane wavelengths (range)	1272.55 to 1274.54 1276.89 to 1278.89 1281.25 to 1283.27 1285.65 to 1287.68 1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19		nm
Side-mode suppression ratio (SMSR), (min)	30		dB
Total average launch power (max)	13.2		dBm
Average launch power, each lane [a] (max)	5.3		dBm
Average launch power, each lane [b] (min)	-3.5	-2.8	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (max)	5.5	5.7	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (min) [c]	-0.5	0.2	dBm
Difference in launch power between any two lanes (OMA_{outer}) (max)	4		dB
Launch power in OMA_{outer} minus TDECQ, each lane (min):			

for extinction ratio ≥ 4.5dB for extinction ratio < 4.5dB	-1.9 -1.8	-1.2 -1.1	dBm dBm
Transmitter and dispersion eye closure for PAM4 (TDECQ), each lane (max)	3.1	3.3	dB
Average launch power of OFF transmitter, each lane (max)	-30		dBm
Extinction ratio, each lane (min)	3.5		dB
$RIN_{16.5}OMA$ (max)	-132	-	dB/Hz
$RIN_{15.1}OMA$ (max)	-	-132	dB/Hz
Optical return loss tolerance (max)	16.5	15.1	dB
Transmitter reflectance [d] (max)	-26		dB

[a] As the total average launch power limit has to be met, not all of the lanes can operate at the maximum average launch power, each lane.

[b] Average launch power, each lane (min) is informative and not the principal indicator of signal strength. A transmitter with launch power below this value cannot be compliant; however, a value above this does not ensure compliance.

[c] Even if the TDECQ < 1.4dB for an extinction ratio of ≥ 4.5dB or TDECQ < 1.3dB for an extinction ratio of < 4.5dB, the OMA_{outer} (min) must exceed this value.

[d] Transmitter reflectance is defined looking into the transmitter.

Table 3-27. *400GBASE-FR8 and 400GBASE-LR8 transmit characteristics.*

3.7.3.2 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 receive optical specifications

The 200GBASE-FR4 receiver shall meet the specifications defined in **Table 3-28** per the definitions in Section **3.7.4**. The 200GBASE-LR4 receiver shall meet the specifications defined in **Table 3-28** per the definitions in Section **3.7.4**. The 400GBASE-FR8 receiver shall meet the specifications defined in **Table 3-29** per the definitions in Section **3.7.4**. The 400GBASE-LR8 receiver shall meet the specifications defined in **Table 3-29** per the definitions in Section **3.7.4**.

Description	200GBASE-FR4	200GBASE-LR4	Unit
Signaling rate, each lane (range)	26.5625 ± 100 ppm		GBd
Modulation format	PAM4		-
Lane wavelengths (range)	1264.5 to 1277.5 1284.5 to 1297.5 1304.5 to 1317.5 1324.5 to 1337.5	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19	nm

Damage threshold [a], each lane	5.7	6.3	dBm
Average receive power, each lane (max)	4.7	5.3	dBm
Average receive power, each lane [b] (min)	-8.2	-9.7	dBm
Receive power (OMA_{outer}), each lane (max)	4.5	5.1	dBm
Difference in receive power between any two lanes (OMA_{outer}) (max)	4.1	4.2	dB
Receiver reflectance (max)	-26		dB
Receiver sensitivity (OMA_{outer}), each lane [c] (max)	-6	-7.7	dBm
Stressed receiver sensitivity (OMA_{outer}), each lane [d] (max)	-3.6	-5.2	dBm
Conditions of stressed receiver sensitivity test: [e]			
Stressed eye closure for PAM4 (SECQ), lane under test	3.3	3.4	dB
OMA_{outer} of each aggressor lane	0.5	-1	dBm

[a] The receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level.

[b] Average receive power, each lane (min) is informative and not the principal indicator of signal strength. A received power below this value cannot be compliant; however, a value above this does not ensure compliance.

[c] Receiver sensitivity (OMA_{outer}), each lane (max) is informative and is defined for a transmitter with SECQ of 0.9dB.

[d] Measured with conformance test signal at TP3 (see Section **3.7.4.13**) for the BER specified in Section **3.7**.

[e] These test conditions are for measuring stressed receiver sensitivity. They are not characteristics of the receiver.

Table 3-28. *200GBASE-FR4 and 200GBASE-LR4 receive characteristics.*

Description	400GBASE-FR8	400GBASE-LR8	Unit
Signaling rate, each lane (range)	26.5625 ± 100 ppm		GBd
Modulation format	PAM4		-
Lane wavelengths (range)	1272.55 to 1274.54 1276.89 to 1278.89 1281.25 to 1283.27 1285.65 to 1287.68		nm

	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19		
Damage threshold [a], each lane	6.3		dBm
Average receive power, each lane (max)	5.3		dBm
Average receive power, each lane [b] (min)	-7.5	-9.1	dBm
Receive power (OMA_{outer}), each lane (max)	5.7		dBm
Difference in receive power between any two lanes (OMA_{outer}) (max)	4.1	4.5	dB
Receiver reflectance (max)	-26		dB
Receiver sensitivity (OMA_{outer}), each lane [c] (max)	-5.3	-7.1	dBm
Stressed receiver sensitivity (OMA_{outer}), each lane [d] (max)	-3.1	-4.7	dBm
Conditions of stressed receiver sensitivity test: [e]			
Stressed eye closure for PAM4 (SECQ), lane under test	3.1	3.3	dB
OMA_{outer} of each aggressor lane	1	-0.2	dBm

[a] The receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level.

[b] Average receive power, each lane (min) is informative and not the principal indicator of signal strength. A received power below this value cannot be compliant; however, a value above this does not ensure compliance.

[c] Receiver sensitivity (OMA_{outer}), each lane (max) is informative and is defined for a transmitter with SECQ of 0.9dB.

[d] Measured with conformance test signal at TP3 (see Section **3.7.4.13**) for the BER specified in Section **3.7**.

[e] These test conditions are for measuring stressed receiver sensitivity. They are not characteristics of the receiver.

Table 3-29. *400GBASE-FR8 and 400GBASE-LR8 receive characteristics.*

3.7.3.3 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 illustrative link power budgets

Illustrative power budgets and penalties for 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 channels are shown in **Table 3-30**.

Parameter	200GBASE-FR4	400GBASE-FR8	200GBASE-LR4	400GBASE-LR8	Unit
Power budget (for maximum TDECQ):					
for extinction ratio ≥ 4.5dB	7.6	7.4	10.2	10.1	dB
for extinction ratio < 4.5dB	7.7	7.5	10.3	10.2	dB
Operating distance	2		10		km
Channel insertion loss	4 [a]		6.3		dB
Maximum discrete reflectance	See Section **3.7.7**		See Section **3.7.7**		dB
Allocation for penalties [b] (for maximum TDECQ):					
for extinction ratio ≥ 4.5dB	3.6	3.4	3.9	3.8	dB
for extinction ratio < 4.5dB	3.7	3.5	4	3.9	dB
Additional insertion loss allowed	0		0		dB

[a] The channel insertion loss is calculated using the maximum distance specified in **Table 3-25** for 200GBASE-FR4 and 400GBASE-FR8 and fiber attenuation of 0.4dB/km plus an allocation for connection and splice loss given in Section **3.7.7**.

[b] Link penalties are used for link budget calculations. They are not requirements and are not meant to be tested.

Table 3-30. *200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 illustrative link power budgets.*

3.7.4 Definition of optical parameters and measurements methods

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified.

3.7.4.1 Test patterns for optical parameters

While compliance is to be achieved in normal operation, specific test patterns are defined for measurement consistency and to enable measurement of some

parameters. **Table 3-32** gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the sections in which each parameter is defined. Any of the test patterns given for a particular test in **Table 3-32** may be used to perform that test. The test patterns used in this Section are shown in **Table 3-31**.

Pattern	Pattern description
Square wave	Square wave (8 threes, 8 zeros)
3	PRBS31Q
4	PRBS13Q
5	Scrambled idle
6	SSPRQ

Table 3-31. *Test patterns.*

Parameter	Pattern	Related section
Wavelength	Square wave, 3, 4, 5, 6 or valid 200GBASE-R or 400GBASE-R signal	**3.7.4.2**
Side mode suppression ratio	3, 5, 6 or valid 200GBASE-R or 400GBASE-R signal	**-**
Average optical power	3, 5, 6 or valid 200GBASE-R or 400GBASE-R signal	**3.7.4.3**
Outer Optical Modulation Amplitude (OMA_{outer})	4 or 6	**3.7.4.4**
Transmitter and dispersion eye closure for PAM4 (TDECQ)	6	**3.7.4.5**
Extinction ratio	4 or 6	**3.7.4.10**
$RIN_{15.1}OMA$ and $RIN_{16.5}OMA$	Square wave	**3.7.4.11**
Stressed receiver conformance test signal calibration	6	**3.7.4.13**
Stressed receiver sensitivity	3 or 5	**3.7.4.13**

Table 3-32. *Test-pattern definitions and related sections.*

3.7.4.2 Wavelength

The wavelength of each optical lane shall be within the ranges given in **Table 3-26** for 200GBASE-FR4, in **Table 3-23** for 200GBASE-LR4 and **Table 3-24** for

400GBASE-FR8 and 400GBASE-LR8, if measured per IEC 61280-1-3. The lane under test is modulated using the test pattern defined in **Table 3-32**.

3.7.4.3 Average optical power

The average optical power of each lane shall be within the limits given in **Table 3-26** for 200GBASE-FR4 and 200GBASE-LR4 and in **Table 3-27** for 400GBASE-FR8 and 400GBASE-LR8 if measured using the methods given in IEC 61280-1-1, with the sum of the optical power from all of the lanes not under test below –30dBm, per the test setup in **Figure 3-17**.

3.7.4.4 Outer Optical Modulation Amplitude (OMA_{outer})

The OMA_{outer} of each lane shall be within the limits given in **Table 3-26** for 200GBASE-FR4 and 200GBASE-LR4 and in **Table 3-27** for 400GBASE-FR8 and 400GBASE-LR8. The OMA_{outer} is measured using a test pattern specified for OMA_{outer} in **Table 3-32**. It is the difference between the average optical launch power level P_3, measured over the central 2 UI of a run of 7 threes, and the average optical launch power level P_0, measured over the central 2 UI of a run of 6 zeros, as shown in **Figure 3-27**. For this measurement the sum of the optical power from all of the lanes not under test is below –30dBm, or if other lanes are operating, a suitable optical filter may be used to separate the lane under test.

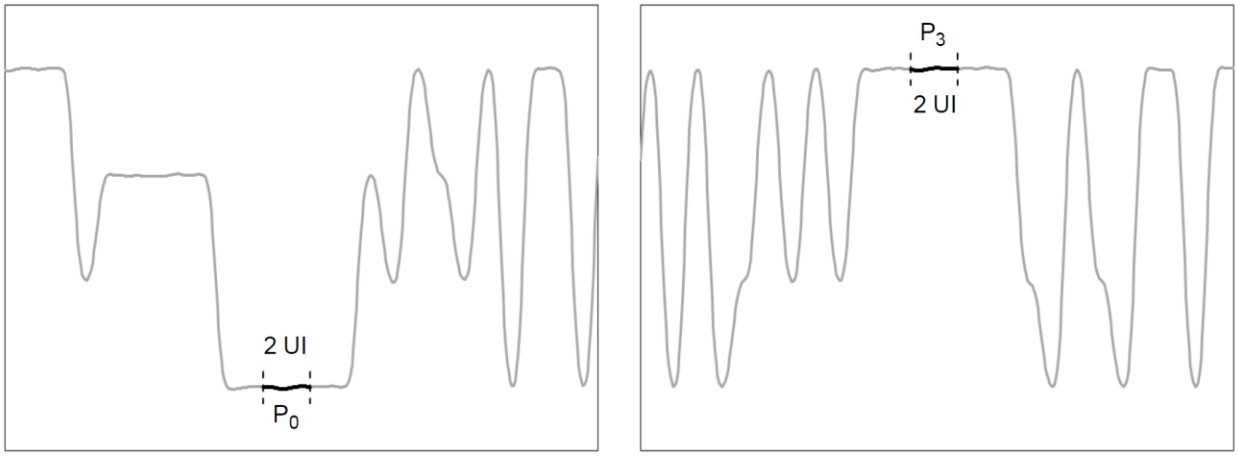


Figure 3-27. *Power levels P_0 and P_3 from PRBS13Q test pattern.*

3.7.4.5 Transmitter and dispersion eye closure for PAM4 (TDECQ)

The TDECQ of each lane shall be within the limits given in **Table 3-26** for 200GBASE-FR4 and 200GBASE-LR4 and in **Table 3-27** for 400GBASE-FR8 and

400GBASE-LR8 if measured using the methods specified in Sections **3.7.4.6**, **3.7.4.7**, and **3.7.4.8**.

TDECQ is a measure of each optical transmitter's vertical eye closure when transmitted through a worst-case optical channel (specified in Section **3.7.4.7**), as measured through an optical to electrical converter (O/E) with a bandwidth equivalent to a reference receiver and equalized with the reference equalizer (as described in Section **3.7.4.9**). The reference receiver and equalizer may be implemented in software or may be part of the oscilloscope.

Table 3-32 specifies the test patterns to be used for measurement of TDECQ.

3.7.4.6 TDECQ conformance test setup

A block diagram for the TDECQ conformance test is shown in **Figure 3-28**. Other equivalent measurement implementations may be used with suitable calibration.

Each optical lane is tested individually with all other lanes in operation and all lanes using the same test pattern. There shall be at least 31 UI delay between the test pattern on one lane and the pattern on any other lane, so that the symbols on each lane are not correlated within the PMD. The optical splitter and variable reflector are adjusted so that each transmitter is tested with the optical return loss specified in **Table 3-33**. The state of polarization of the back reflection is adjusted to create the greatest RIN. The optical filter is used to separate the lane under test from the others. Each optical lane is tested with the optical channel described in Section **3.7.4.7**. The combination of the O/E and the oscilloscope has a fourth-order Bessel-Thomson filter response with a bandwidth of approximately 13.28125GHz. Compensation may be made for any deviation from an ideal fourth-order Bessel-Thomson response.

The optical filter passband ripple shall be limited to 0.5dB peak-to-peak and the isolation is chosen such that the ratio of the power in the lane being measured to the sum of the powers of all of the other lanes is greater than 20dB (see ITU-T G.959.1 Annex B).

The test pattern (specified in **Table 3-32**) is transmitted repetitively by the optical lane under test and the oscilloscope is set up to capture the complete pattern for TDECQ analysis as described in Section **3.7.4.8**. The clock recovery unit (CRU) has a corner frequency of 4MHz and a slope of 20dB/decade. The CRU can be implemented in hardware or software depending on oscilloscope technology.

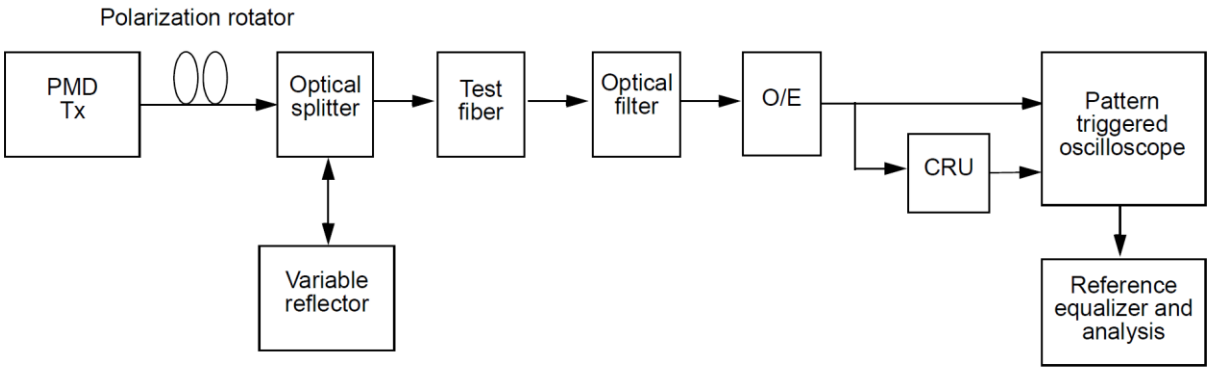


Figure 3-28. *TDECQ conformance test block diagram.*

3.7.4.7 Channel requirements

The transmitter is tested using an optical channel that meets the requirements listed in **Table 3-33**.

PMD type	Dispersion [a] (ps/nm)		Insertion loss [b]	Optical return loss [c]	Max mean DGD
	Minimum	Maximum			
200GBASE-FR4 or 400GBASE-FR8	$0.0465 \cdot \lambda \cdot [1 - (1324/\lambda)^4]$	$0.0465 \cdot \lambda \cdot [1 - (1300/\lambda)^4]$	Minimum	17.8dB	0.8ps
200GBASE-LR4 or 400GBASE-LR8	$0.2325 \cdot \lambda \cdot [1 - (1324/\lambda)^4]$	$0.2325 \cdot \lambda \cdot [1 - (1300/\lambda)^4]$	Minimum	15.7dB	0.8ps

[a] The dispersion is measured for the wavelength of the device under test (λ in nm). The coefficient assumes 2km for 200GBASE-FR4 and 400GBASE-FR8 and 10km for 200GBASE-LR4 and 400GBASE-LR8.

[b] There is no intent to stress the sensitivity of the O/E converter associated with the oscilloscope.

[c] The optical return loss is applied at TP2.

Table 3-33. *Transmitter compliance channel specifications.*

A 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8 or 400GBASE-LR8 transmitter is to be compliant with a total dispersion at least as negative as the "minimum dispersion" and at least as positive as the "maximum dispersion" columns specified in **Table 3-33** for the wavelength of the device under test. This may be achieved with channels consisting of fibers with lengths chosen to meet the dispersion requirements.

To verify that the fiber has the correct amount of dispersion, the measurement method defined in IEC 60793-1-42 may be used. The measurement is made in the linear power regime of the fiber.

The channel provides an optical return loss specified in **Table 3-33**. The state of polarization of the back reflection is adjusted to create the greatest RIN.

The mean DGD of the channel is to be less than the value specified in **Table 3-33**.

3.7.4.8 TDECQ measurement method

TDECQ for 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 is measured as described in Section **3.6.4.8** with the exception that the reference equalizer is as specified in Section **3.7.4.9**.

3.7.4.9 TDECQ reference equalizer

The reference equalizer for 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 is a 5 tap, *T* spaced, Feed-Forward Equalizer (FFE), where *T* is the symbol period. The sum of the equalizer tap coefficients is equal to 1.

NOTE: This reference equalizer is part of the TDECQ test and does not imply any particular receiver equalizer implementation.

3.7.4.10 Extinction ratio

The extinction ratio of each lane shall be within the limits given in **Table 3-26** for 200GBASE-FR4 and 200GBASE-LR4 and in **Table 3-27** for 400GBASE-FR8 and 400GBASE-LR8 if measured using a test pattern specified for extinction ratio in **Table 3-32** with the sum of the optical power from all of the lanes not under test being below –30dBm, or if other lanes are operating, a suitable optical filter may be used to separate the lane under test. The extinction ratio of a PAM4 optical signal is defined as the ratio of the average optical launch power level P_3, measured over the central 2 UI of a run of 7 threes, and the average optical launch power level P_0, measured over the central 2 UI of a run of 6 zeros, as shown in **Figure 3-27**.

3.7.4.11 Relative intensity noise ($RIN_{16.5}OMA$ and $RIN_{15.1}OMA$)

RIN shall be as defined by the measurement methodology of Relative Intensity Noise Optical Modulation Amplitude (RIN_xOMA) with the following exceptions:

a. The optical return loss is 16.5dB for 200GBASE-FR4 and 400GBASE-FR8 and 15.1dB for 200GBASE-LR4 and 400GBASE-LR8.
b. Each lane may be tested individually with the sum of the optical power from all of the lanes not under test being below -30dBm, or if other lanes are operating, a suitable optical filter may be used to separate the lane under test.
c. The upper -3dB limit of the measurement apparatus is to be approximately equal to the signaling rate (i.e., 26.6GHz).
d. The test pattern is according to **Table 3-32**.

3.7.4.12 Receiver sensitivity

Receiver sensitivity, which is defined for an input signal with SECQ of 0.9dB (e.g., an ideal input signal without overshoot), is informative and compliance is not required. If measured, the test signal should have negligible impairments such as ISI, rise/fall times, jitter and RIN. Instead, the normative requirement for receivers is stressed receiver sensitivity.

3.7.4.13 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in **Table 3-28** for 200GBASE-FR4 and 200GBASE-LR4 and in **Table 3-29** for 400GBASE-FR8 and 400GBASE-LR8, using the test pattern specified for SRS in **Table 3-32**. The BER is required to be met for the lane under test on its own.

Stressed receiver sensitivity is defined with all transmit and receive lanes in operation. Any of the patterns specified for SRS in **Table 3-32** is sent from the transmit section of the PMD under test. The signal being transmitted is asynchronous to the received signal.

3.7.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 optical transceivers shall conform to Hazard Level 1 laser requirements as defined in IEC 60825-1 and IEC 60825-2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Conformance to additional laser safety standards may be required for operation within specific geographic regions.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service. This documentation explicitly defines requirements and usage restrictions on the host system necessary to meet these safety certifications.[2]

It is recommended that proper installation practices, as defined by applicable local codes and regulation, be followed in every instance in which such practices are applicable.

[2] A host system that fails to meet the manufacturer's requirements and/or usage restrictions may emit laser radiation in excess of the safety limits of one or more safety standards. In such a case, the host manufacturer is required to obtain its own laser safety certification.

Normative specifications in this Section shall be met by a system integrating a 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, or 400GBASE-LR8 PMD over the life of the product while the product operates within the manufacturer's range of environmental, power, and other specifications.

It is recommended that manufacturers indicate in the literature associated with the PHY the operating environmental conditions to facilitate selection, installation, and maintenance. It is recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of this Section will be met.

A system integrating a 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, or 400GBASE-LR8 PMD shall comply with applicable local and national codes for the limitation of electromagnetic interference.

The optical link is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration).

3.7.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 3-29**.

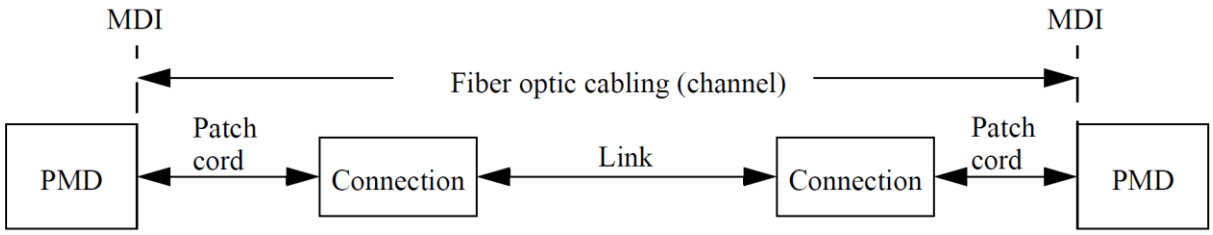


Figure 3-29. *Fiber optic cabling model.*

The channel insertion loss is given in **Table 3-34**. A channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion meet the specifications. Insertion loss measurements of installed fiber cables are made in accordance with IEC 6128-1-1. The fiber optic cabling model (channel) defined here is the same as a simplex fiber optic link segment. The term channel is used here for consistency with generic cabling standards.

Description	200GBASE-FR4	400GBASE-FR8	200GBASE-LR4	400GBASE-LR8	Unit
Operating distance (max)	2		10		km
Channel insertion loss [a, b] (max)	4		6.3 [c]		dB
Channel insertion loss (min)	0		0		dB
Positive dispersion [b] (max)	6.7	1.9	9.5		ps/nm
Negative dispersion [b] (min)	-11.9	-10.2	-28.4	-50.8	ps/nm
DGD_max [d]	3		8		ps
Optical return loss (min)	25		22		dB

[a] These channel insertion loss values include cable, connectors, and splices.

[b] Over the wavelength range 1264.5nm to 1337.5nm for 200GBASE-FR4, 1294.53nm to 1310.19nm for 200GBASE-LR4, and 1272.55nm to 1310.19nm for 400GBASE-FR8 and 400GBASE-LR8.

[c] Using 0.46dB/km at 1272.55nm attenuation for optical fiber cables derived from Appendix I of ITU-T G.695 may not support operation at 10km for 400GBASE-LR8 under worst case conditions.

[d] Differential Group Delay (DGD) is the time difference at reception between the fractions of a pulse that were transmitted in the two principal states of polarization of an optical signal. *DGD_max* is the maximum differential group delay that the system must tolerate.

Table 3-34. *Fiber optic cabling (channel) characteristics.*

3.7.7 Characteristics of the fiber optic cabling (channel)

The 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, and 400GBASE-LR8 fiber optic cabling shall meet the specifications defined in **Table 3-34**. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

The fiber optic cable requirements are satisfied by cables containing IEC 60793-2-50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-mode), or type B6_a (bend insensitive) fibers or the requirements in **Table 3-35** where they differ.

Description	Value	Unit
Nominal fiber specification wavelength	1310	nm
Cabled optical fiber attenuation (max)	0.47 [a] or 0.5	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0.093	ps/nm^2km
[a] The 0.47dB/km at 1264.5nm attenuation for optical fiber cables is derived from Appendix I of ITU-T G.695.		

Table 3-35. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 3-29**, consists of a mated pair of optical connectors.

The maximum link distance for 200GBASE-LR4 and 400GBASE-LR8 is based on an allocation of 2dB total connection and splice loss. For example, this allocation supports four connections with an average insertion loss per connection of 0.5dB. The maximum link distance for 200GBASE-FR4 and 400GBASE-FR8 is based on an allocation of 3dB total connection and splice loss. Connections with different loss characteristics may be used provided the requirements of **Table 3-34** are met.

The maximum value for each discrete reflectance shall be less than or equal to the value shown in **Table 3-36** corresponding to the number of discrete reflectances above –55dB within the channel. For numbers of discrete reflectances in between two numbers shown in the table, the lower of the two corresponding maximum discrete reflectance values applies.

Number of discrete reflectances above –55 dB	Maximum value for each discrete reflectance	
	200GBASE-FR4 or 400GBASE-FR8	200GBASE-LR4 or 400GBASE-LR8
1	-25dB	-22dB
2	-31dB	-29dB
4	-35dB	-33dB
6	-38dB	-35dB
8	-40dB	-37dB
10	-41dB	-39dB

Table 3-36. *Maximum value of each discrete reflectance.*

3.7.7.1 MDI requirements

The 200GBASE-FR4, 200GBASE-LR4, 400GBASE-FR8, or 400GBASE-LR8 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the

PMD and the “fiber optic cabling” (as shown in **Figure 3-29**). Examples of an MDI include the following:

a. Connectorized fiber pigtail.

b. PMD connector.

When the MDI is a plug and connector connection, it shall meet the interface performance specifications of IEC 61753-1-1 and IEC 61753-021-2.

NOTE: Transmitter compliance testing is performed at TP2 as defined in Section **3.7.1**, not at the MDI.

3.8 Physical Medium Dependent (PMD) sublayer and medium, type 400GBASE-SR16

400GBASE-SR16 PMD together with the multimode fiber medium. The PMD sublayer provides a point-to-point 400 Gb/s Ethernet link over 16 pairs of multimode fiber, up to at least 100m. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in **Table 3-37**.

Associated section	400GBASE-SR16
3.2 - RS	Required
3.2 - 400GMIIa	Optional
3.3 - 400GMII Extender	Optional
3.4 - PCS	Required
3.5 - PMA	Required
Annex B - Chip-to-chip 400GAUI-16	Optional
Annex C - Chip-to-module 400GAUI-16	Optional
Annex D - Chip-to-chip 400GAUI-8	Optional
Annex E - Chip-to-module 400GAUI-8	Optional
Annex F - Energy Efficient Ethernet	Optional
[a] The 400GMII is an optional interface. However, if the 400GMII is not implemented, a conforming implementation must behave functionally as though the RS and 400GMII were present	

Table 3-37. *Physical Layer sections associated with the 400GBASE-SR16 PMD.*

Figure 3-30 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC OSI reference model. 400 Gb/s Ethernet is introduced in Section **3**.

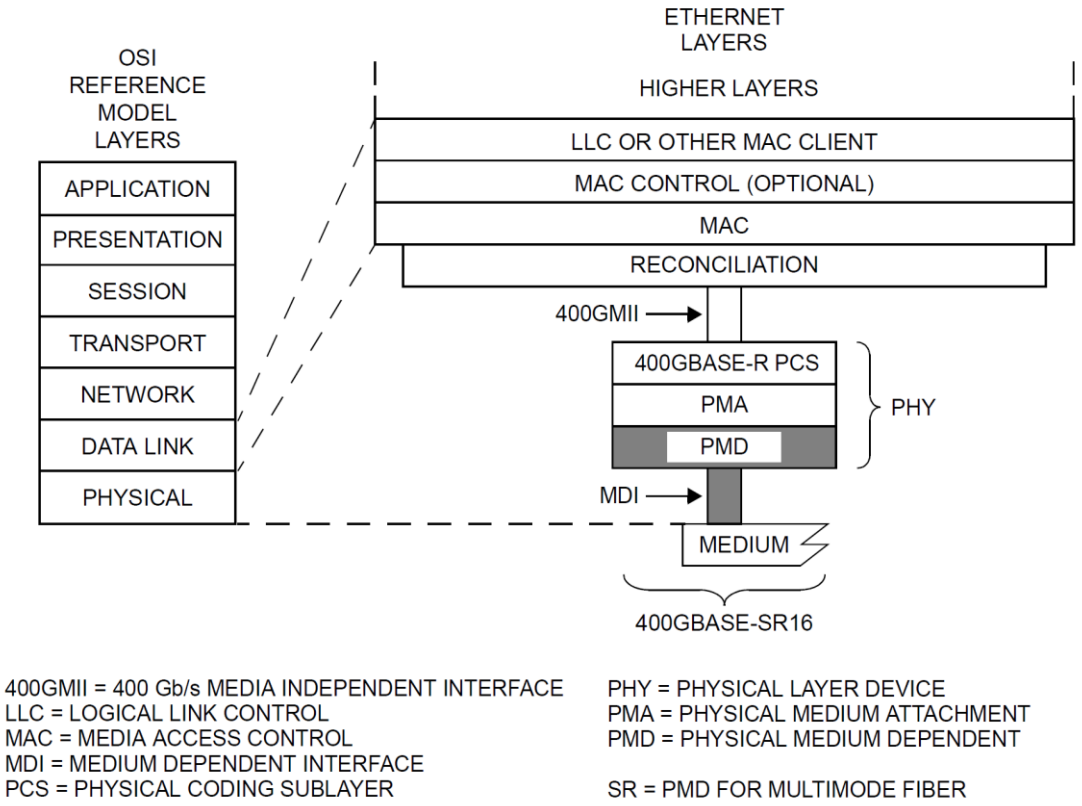


Figure 3-30. *400GBASE-SR16 PMD.*

400GBASE-SR16 PHYs with the optional EEE fast wake capability may enter the LPI mode to conserve energy during periods of low link utilization (see **Annex F**). The deep sleep mode of EEE is not supported.

The BER shall be less than 2.4×10^{-4} provided that the error statistics are sufficiently random that this results in a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap when processed according to Section **3.4**. For a complete PHY, the frame loss ratio may be degraded to 6.2×10^{-11} for 64-octet frames with minimum interpacket gap due to additional errors from the electrical interfaces.

If the error statistics are not sufficiently random to meet this requirement, then the BER shall be less than that required to give a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap.

An upper bound to the delay through the PMA and PMD is required for predictable operation of the MAC Control *PAUSE* operation. The sum of the transmit and receive delays at one end of the link contributed by the 400GBASE-SR16 PMD including 2m of fiber in one direction shall be no more than 8192-bit times (16 *pause_quanta* or 20.48ns). A description of overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **3.1.3** and its references.

The Skew (relative delay) between the lanes must be kept within limits so that the information on the lanes can be reassembled by the PCS sublayer. Skew and Skew Variation are defined in Section **3.1.4** and specified at the points SP1 to SP6 shown in **Figure 3-5** and **Figure 3-6**.

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

3.8.1 PMD Functional Specifications

The 400GBASE-SR16 PMD performs the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

The PMD block diagram is shown in **Figure 3-31**. The optical transmit signal is defined at the output end of a multimode fiber patch cord (TP2), between 2m and 5m in length. Unless specified otherwise, all transmitter measurements and tests defined in Section **3.8.4** are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see Section **3.8.7.1**). Unless specified otherwise, all receiver measurements and tests defined in Section **3.8.4** are made at TP3.

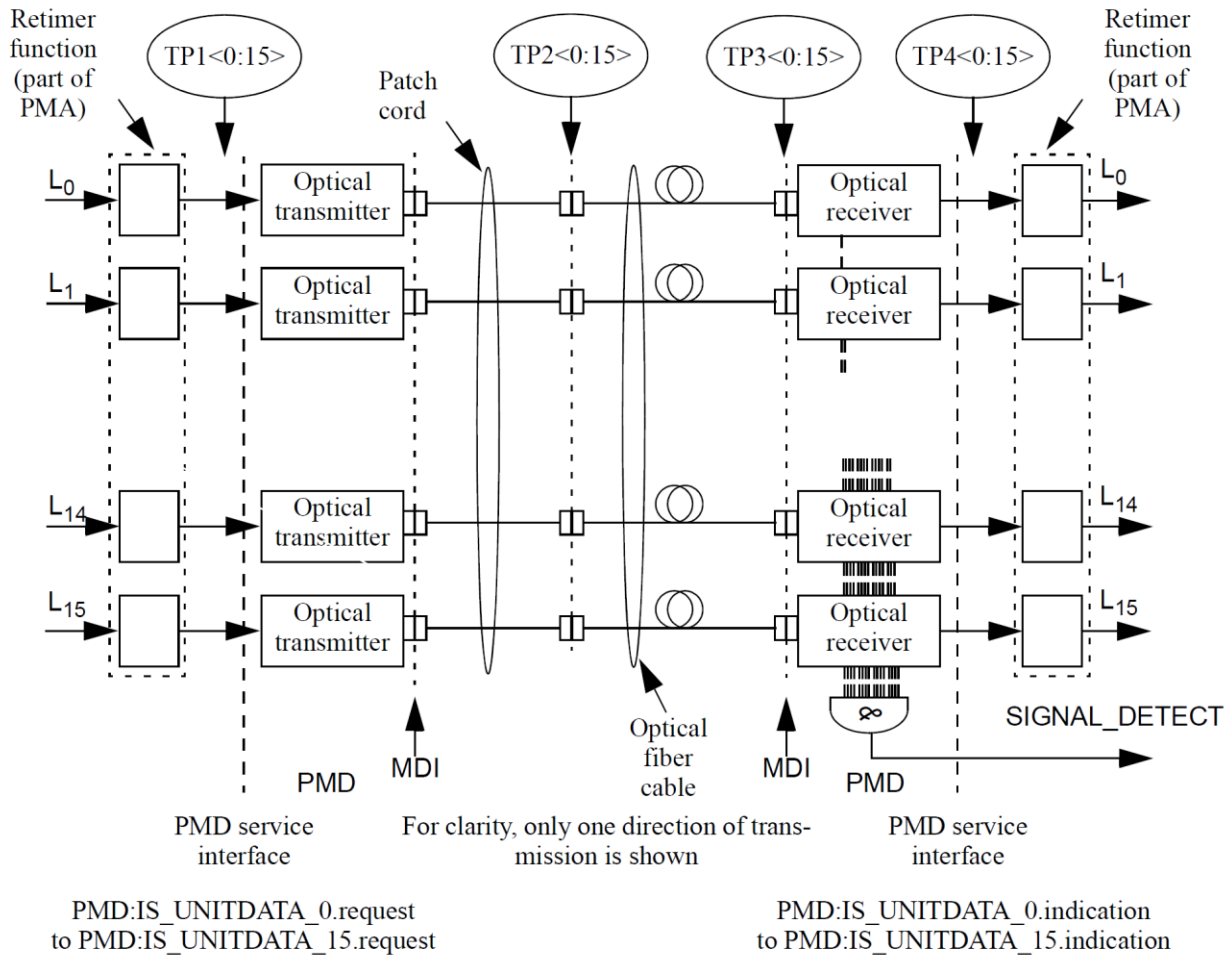


Figure 3-31. *Block diagram for 400GBASE-SR16 transmit/receive paths.*

TP1 and TP4 are informative reference points that may be useful for testing components.

3.8.1.1 PMD transmit function

The PMD Transmit function shall convert the 16 bit streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_15.request* into 16 optical signals. The optical signals shall then be delivered to the MDI which contains 16 parallel light paths for transmit, according to the transmit optical specifications in this section. The higher optical power level in each signal stream shall correspond to *tx_symbol* = *one*.

3.8.1.2 PMD receive function

The PMD Receive function shall convert the 16 optical signals received from the MDI into separate bit streams for delivery to the PMD service interface using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_15.indication*, all

according to the receive optical specifications in this section. The higher optical power level in the signal stream shall correspond to *rx_symbol* = *one*.

3.8.1.3 PMD global signal detect function

The PMD global signal detect function reports the state of the SIGNAL_DETECT via the PMD service interface. The SIGNAL_DETECT parameter is signaled continuously, while the PMD:IS_SIGNAL.indication message is generated when a change occurs in the value of SIGNAL_DETECT. The SIGNAL_DETECT parameter maps to the SIGNAL_OK parameter in the inter-sublayer service interface primitives.

SIGNAL_DETECT is an indicator of the presence of the optical signals on all 16 lanes. The value of the *SIGNAL_DETECT* parameter shall be generated according to the conditions defined in **Table 3-38**. The PMD receiver is not required to verify whether a compliant 400GBASE-SR16 signal is being received. This standard imposes no response time requirements on the generation of the *SIGNAL_DETECT* parameter.

Receive conditions	SIGNAL_DETECT value
For any lane. Average optical power at TP3 ≤ -30dBm	FAIL
For all lanes. [(Optical power at TP3 ≥ average receive power, each lane (min) in 100GBASE-SR4) AND (Compliant 400GBASE-SR16 signal input)]	OK
All other conditions	Unspecified

Table 3-38. *SIGNAL_DETECT value definition.*

All implementations must provide adequate margin between the input optical power level at which the *SIGNAL_DETECT* parameter is set to *OK*, and the inherent noise level of the PMD including the effects of the crosstalk, the power supply noise, etc.

Various implementations of the PMD global signal detect function are permitted by this standard, including implementations that generate the *SIGNAL_DETECT* parameter values in response to the amplitude of the modulation of the optical signals and implementations that respond to the average optical power of the modulated optical signals. When the MDIO is implemented, the *SIGNAL_DETECT* value shall be continuously set in response to the magnitude of the optical signals, according to the requirements of **Table 3-38**.

3.8.1.4 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard. When the MDIO is implemented, each *PMD_signal_detect_i*, where *i* represents the lane number in the range 0:15, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of **Table 3-38**.

3.8.2 Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for 400GBASE-SR16. While it is expected that a PMD will map electrical lane *i* to optical lane *i* and vice versa, there is no need to define the physical ordering of the lanes, as the PCS sublayer is capable of receiving the lanes in any arrangement. The positioning of transmit and receive lanes at the MDI is specified in Section **3.8.7.1**.

3.8.3 PMD to MDI optical specifications for 400GBASE-SR16

The operating range for the 400GBASE-SR16 PMD is defined in **Table 3-39**. A 400GBASE-SR16 compliant PMD operates on 50/125µm multimode fibers, type A1a.2 (OM3), type A1a.3 (OM4), or type A1a.4 (OM5), according to the specifications defined in **Table 3-40**. A PMD that exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., a 400GBASE-SR16 PMD operating at 120m meets the operating range requirement of 0.5m to 100m). The operating range shown in **Table 3-39** is the same as 100GBASE-SR4.

PMD type	Required operating range [a]
400GBASE-SR16	0.5m to 70m for OM3
	0.5m to 100m for OM4
	0.5m to 100m for OM5
[a] The PCS FEC correction function may not be bypassed for any operating distance.	

Table 3-39. *400GBASE-SR16 operating range.*

3.8.3.1 400GBASE-SR16 transmitter optical specifications

The optical characteristics of each lane of a 400GBASE-SR16 transmitter shall be the same as those of a single lane of 100GBASE-SR4, with the exception that the “signaling rate, each lane” is 26.5625GBd ±100ppm, TDEC is as specified in Section **3.8.4.5**, and the BER requirement is as specified in Section **3.8**.

3.8.3.2 400GBASE-SR16 receive optical specifications

The optical characteristics of each lane of a 400GBASE-SR16 receiver shall be the same as those of a single lane of 100GBASE-SR4, with the exception that the "signaling rate, each lane" is 26.5625GBd ±100ppm, stressed receiver sensitivity is as specified in Section **3.8.4.8**, and the BER requirement is as specified in Section **3.8**.

3.8.3.3 400GBASE-SR16 illustrative link power budgets

The illustrative power budget and penalties for each lane of a 400GBASE-SR16 link are the same as those of a single lane of 100GBASE-SR4.

3.8.4 Definition of optical parameters and measurements methods

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified. Optical parameters and measurement methods for 400GBASE-SR16 are the same as for 100GBASE-SR4.

3.8.4.1 Test patterns for optical parameters

The test patterns used in this Section shall be the same as those used for 100GBASE-SR4 with the exception that pattern 5 defined in PCS for 400GBASE-SR16.

3.8.4.2 Center wavelength and spectral width

The center wavelength and RMS spectral width shall be within the range given in 100GBASE-SR4 if measured per IEC 61280-1-3. The transmitter is modulated using one of the test patterns specified in Section **3.8.4.1**.

3.8.4.3 Average optical power

The average optical power shall be within the limits given in 100GBASE-SR4 if measured using the methods given in IEC 61280-1-1. The average optical power is measured using one of the test patterns specified in Section **3.8.4.1**.

3.8.4.4 Optical Modulation Amplitude (OMA)

OMA shall be within the limits given in 100GBASE-SR4 if measured with a square wave (8 ones, 8 zeros) test pattern or with a PRBS9 test pattern. See Section **3.8.4.1** for test pattern information.

3.8.4.5 Transmitter and dispersion eye closure (TDEC)

TDEC is a measure of the optical transmitter's vertical eye closure. TDEC shall be within the limits given in 100GBASE-SR4 if measured using the methods specified in 10GBASE-SR4, with the exceptions that the clock recovery unit (CRU) has a corner

frequency of 4MHz and in **Equation 2-13**, 3.8906R is replaced by 3.4917R, for consistency with the BER of 2.4×10^{-4} given in Section **3.8**.

3.8.4.6 Extinction ratio

The extinction ratio shall be within the limits given in 100GBASE-SR4 if measured using the methods specified in IEC 61280-2-2. The extinction ratio is measured using one of the test patterns specified for extinction ratio in 100GBASE-SR4.

3.8.4.7 Transmitter optical waveform (transmit eye)

The transmitter optical waveform of a port transmitting the test pattern specified in 100GBASE-SR4 as modified by Section **3.8.4.1** shall meet the transmitter eye mask coordinates and hit ratio specified in **Table 2-39** if measured according to the methods specified in Section 2.6.4.7 with the exception that the clock recovery unit's high frequency corner bandwidth is 4MHz.

3.8.4.8 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in 100GBASE-SR4 if measured using the methods defined in 100GBASE-SR4 with the following exceptions:

- The signaling rate of the test pattern generator is set to the rate defined in Section **3.8.3.2**.
- The Clock Recovery Unit (CRU) has a corner bandwidth of 4MHz.
- Sinusoidal jitter is added as specified in 100GBASE-SR4.
- When using 100GBASE-SR4 to measure the SEC of the stressed receiver conformance test signal, the BER requirement is as specified in Section **3.8** and 3.4917R replaces 3.8906R in **Equation 2-13**.
- The hit ratio for the stressed receiver eye mask definition is changed from 5×10^{-5} to 2.4×10^{-4}.

3.8.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

400GBASE-SR16 optical transceivers shall conform to Hazard Level 1M laser requirements as defined in IEC 60825-1 and IEC 60825-2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Conformance to additional laser safety standards may be required for operation within specific geographic regions.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service.

It is recommended that proper installation practices, as defined by applicable local codes and regulation, be followed in every instance in which such practices are applicable.

Normative specifications in this Section shall be met by a system integrating a 400GBASE-SR16 PMD over the life of the product while the product operates within the manufacturer's range of environmental, power, and other specifications.

It is recommended that manufacturers indicate, in the literature associated with the PHY, the operating environmental conditions to facilitate selection, installation, and maintenance. It is recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of this Section will be met.

A system integrating a 400GBASE-SR16 PMD shall comply with applicable local and national codes for the limitation of electromagnetic interference.

3.8.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 3-32**. The fiber type and length are the same as 100BASE-SR4.

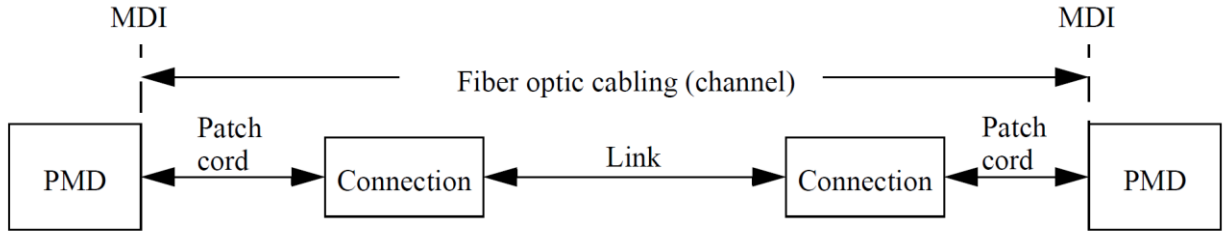


Figure 3-32. *Fiber optic cabling model.*

The channel insertion loss is given in **Table 3-40**.Table **3-40.** Fiber optic cabling (channel) characteristics for 400GBASE-SR16.

A channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, modal dispersion, reflections and losses of all connectors and splices meet the specifications. Insertion loss measurements of installed fiber cables are made in accordance with IEC 61280-4-1. As OM4 and OM5 optical fiber meet the requirements for OM3, a channel compliant to the "OM3" column may use OM4 or OM5 optical fiber, or a combination of OM3, OM4 and OM5. The fiber

optic cabling model (channel) determined here is the same as a simplex fiber optic permanent link. The term channel is used here for consistency with generic cabling standards.

Description	OM3	OM4	OM5	Unit
Operating distance (max)	700	100	100	m
Channel insertion loss [a] (max)	1.8	1.9	1.9	dB
Channel insertion loss (min)	0			dB
[a] These channel insertion loss values include cable loss plus 1.5dB allocated for connection and splice loss, over the wavelength range 840nm to 860nm.				

Table 3-40. *Fiber optic cabling (channel) characteristics for 400GBASE-SR16.*

3.8.7 Characteristics of the fiber optic cabling (channel)

The 400GBASE-SR16 fiber optic cabling shall meet the specifications defined in **Table 3-40**. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together. The requirements for the optical fiber, connection insertion loss, and maximum discrete reflectance are the same as 100GBASE-SR4.

The fiber contained within the 400GBASE-SR16 fiber optic cabling shall comply with the specifications and parameters of **Table 3-41**. A variety of multimode cable types may satisfy these requirements, provided the resulting channel also meets the specifications of **Table 3-40**.

Description	OM3 [a]	OM4 [b]	OM5 [c]	Unit
Nominal core diameter	50			µm
Nominal fiber specification wavelength	850			nm
Effective modal bandwidth (min) [d]	2000	4700		MHz/km
Cabled optical fiber attenuation (max)	3.5			dB/km
Cabled optical fiber attenuation (max)	$1295 \le \lambda_0 \le 1340$		$1297 \le \lambda_0 \le 1328$	nm
Chromatic dispersion slope (max) (S_0)	0.105 for $1295 \le \lambda_0 \le 1310$ and $0.000375 \times (1590 - \lambda_0)$ for $1310 \le \lambda 0 \le 1340$		$-412/(840(1-(\lambda_0/840)^4))$ for $1297 \le \lambda 0 \le 1328$	ps/nm^2km

[a] IEC 60793-2-10 type A1a.2

[b] IEC 60793-2-10 type A1a.3

[c] IEC 60793-2-10 type A1a.4

[d] When measured with the launch conditions specified in 100GBASE-SR4.

Table 3-41. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 3-32**, consists of a mated pair of optical connectors.

The maximum link distance is based on an allocation of 1.5dB total connection and splice loss. For example, this allocation supports 3 connections with an average insertion loss per connection of 0.5dB. Connections with lower loss characteristics may be used provided the requirements of **Table 3-40** are met. However, the loss of a single connection shall not exceed 0.75dB.

The maximum discrete reflectance shall be less than –20dB.

3.8.7.1 Medium Dependent Interface (MDI)

The 400GBASE-SR16 PMD is coupled to the fiber optic cabling at the MDI, which is the interface between the PMD and the “fiber optic cabling” (as shown in **Figure 3-32**). Examples of an MDI include the following:

a. Connectorized fiber pigtail, and
b. PMD connector.

3.8.7.2 Optical lane assignments

The 16 transmit and 16 receive optical lanes of 400GBASE-SR16 shall occupy the positions depicted in **Figure 3-33** when looking into the MDI connector with the connector keyway feature on top. The interface contains 32 active lanes. The transmit optical lanes occupy the upper 16 positions. The receive optical lanes occupy the lower 16 positions.

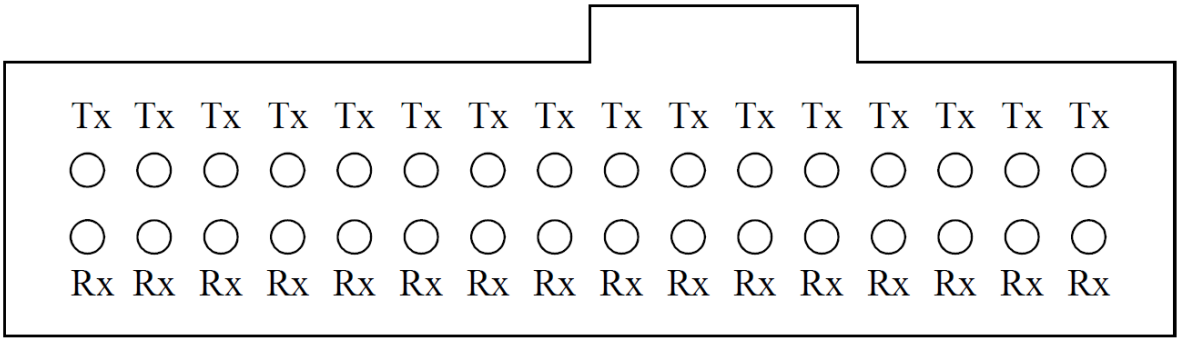


Figure 3-33. *400GBASE-SR16 optical lane assignments.*

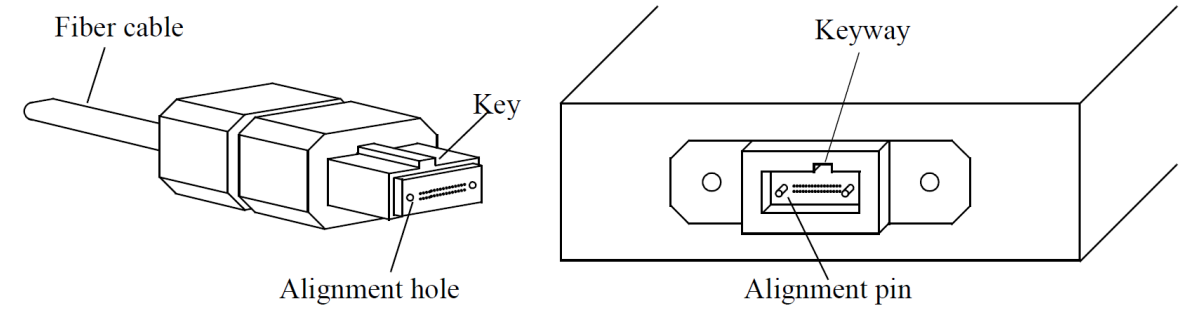


Figure 3-34. *MPO-16 female two-row plug with flat interface, and an MDI.*

3.8.7.3 MDI requirements

The MDI adapter or connector shall meet the dimensional specifications as defined in IEC 61754-7-1. The plug terminating the optical fiber cabling shall meet the dimensional specifications as defined in IEC 61754-7-1. The MDI shall optically mate with the plug on the optical fiber cabling. **Figure 3-34** shows an MPO-16 female two-row plug with flat interface, and an MDI. The MDI shall meet the interface performance specifications of IEC 61753-1 and IEC 61753-022-2 for performance Class Cm/2m.

NOTE: Transmitter compliance testing is performed at TP2 as defined in Section **3.8.1**, not at the MDI.

3.9 Physical Medium Dependent (PMD) sublayer and medium, type 400GBASE-DR4

400GBASE-DR4 PMD is specified together with the single-mode fiber medium. The optical signal generated by this PMD type is modulated using a 4-level pulse amplitude modulation (PAM4) format. When forming a complete PHY, a PMD shall be connected to the appropriate PMA as shown in **Table 3-42**, to the medium through the MDI and optionally with the management functions that may be accessible through the management interface (MDIO), or equivalent.

Associated section	400GBASE-DR4
3.2 - RS	Required
3.2 - 400GMII [a]	Optional
3.3 - 400GMII Extender	Optional
3.4 - PCS	Required
3.5 - PMA	Required
Annex B - Chip-to-chip 400GAUI-16	Optional

Annex C - Chip-to-module 400GAUI-16	Optional
Annex D - Chip-to-chip 400GAUI-8	Optional
Annex E - Chip-to-module 400GAUI-8	Optional
Annex F - Energy Efficient Ethernet	Optional
[a] The 400GMII is an optional interface. However, if the 400GMII is not implemented, a conforming implementation must behave functionally as though the RS and 400GMII were present.	

Table 3-42. *Physical Layer sections associated with the 400GBASE-DR4 PMD.*

Figure 3-35 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC OSI reference model. 400 Gb/s Ethernet is introduced in Section **3** and the purpose of each PHY sublayer is summarized in Section **3.2**.

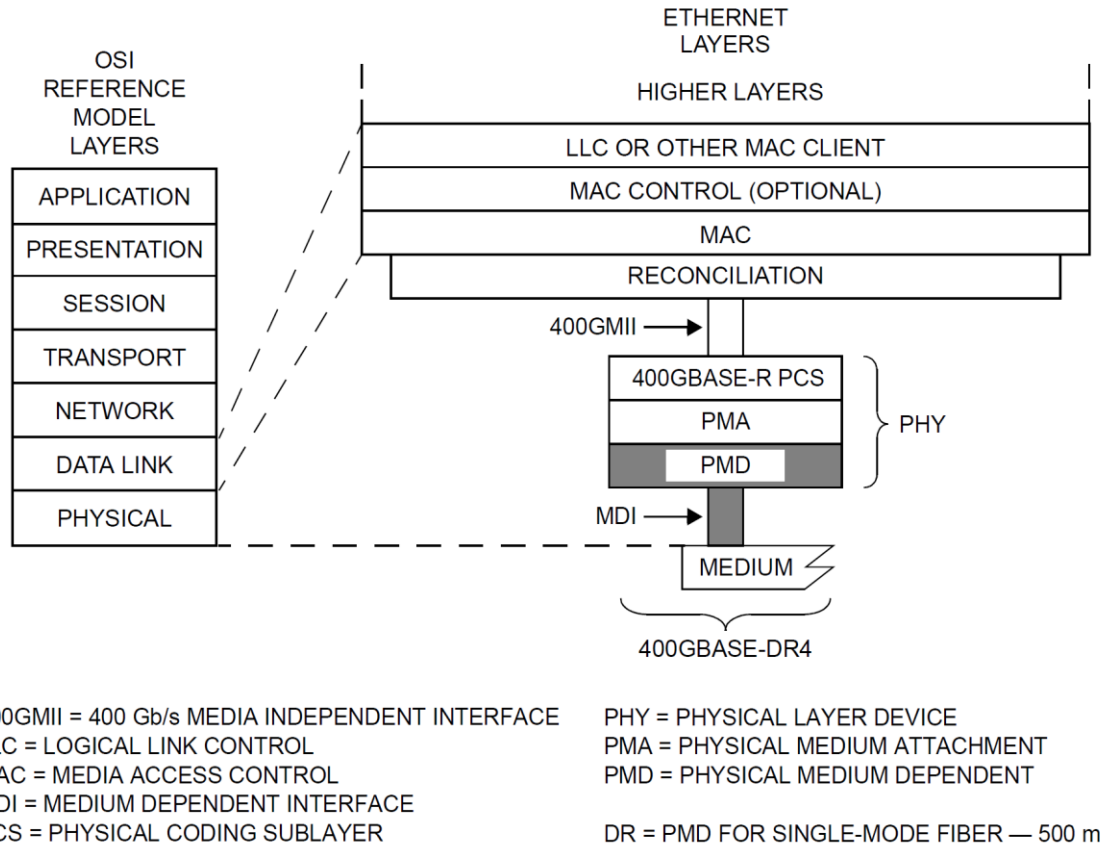


Figure 3-35. *400GBASE-DR4 PMD.*

400GBASE-DR4 PHYs with the optional EEE fast wake capability may enter the Low Power Idle (LPI) mode to conserve energy during periods of low link utilization (see **Annex F**). The deep sleep mode of EEE is not supported.

The BER when processed according to Section **3.5** shall be less than 2.4×10^{-4} provided that the error statistics are sufficiently random that this results in a frame loss

ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap when processed according to Section **3.5** and then Section **3.4**. For a complete PHY, the frame loss ratio may be degraded to 6.2×10^{-11} for 64-octet frames with minimum interpacket gap due to additional errors from the electrical interfaces.

If the error statistics are not sufficiently random to meet this requirement, then the BER shall be less than that required to give a frame loss ratio of less than 1.7×10^{-12} for 64-octet frames with minimum interpacket gap.

The sum of the transmit and receive delays at one end of the link contributed by the 400GBASE-DR4 PMD including 2m of fiber in one direction shall be no more than 8192-bit times (16 *pause_quanta* or 20.48ns). A description of overall system delay constraints and the definitions for bit times and *pause_quanta* can be found in Section **3.1.3** and its references.

The Skew (relative delay between the lanes) and Skew Variation must be kept within limits so that the information on the lanes can be reassembled by the PCS. Skew and Skew Variation are defined in Section **3.1.4** and specified at the points SP1 to SP6 shown in **Figure 3-5** and **Figure 3-6**.

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, then the Skew at SP2 is limited to 43ns and the Skew Variation at SP2 is limited to 400ps.

The Skew at SP3 (the transmitter MDI) shall be less than 54ns and the Skew Variation at SP3 shall be less than 600ps.

The Skew at SP4 (the receiver MDI) shall be less than 134ns and the Skew Variation at SP4 shall be less than 3.4ns.

If the PMD service interface is physically instantiated so that the Skew at SP5 can be measured, then the Skew at SP5 shall be less than 145ns and the Skew Variation at SP5 shall be less than 3.6ns.

3.9.1 PMD Functional Specifications

The 400GBASE-DR4 PMD performs the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

The PMD block diagram is shown in **Figure 3-36**. The optical transmit signal is defined at the output end of a single-mode fiber patch cord (TP2), between 2m and 5m in length. Unless specified otherwise, all transmitter measurements and tests defined in Section **3.9.4** are made at TP2. The optical receive signal is defined at the

output of the fiber optic cabling (TP3) at the MDI (see Section **3.9.7.1**). Unless specified otherwise, all receiver measurements and tests defined in Section **3.9.4** are made at TP3.

TP1<0:3> and TP4<0:3> are informative reference points that may be useful to implementors for testing components (these test points will not typically be accessible in an implemented system).

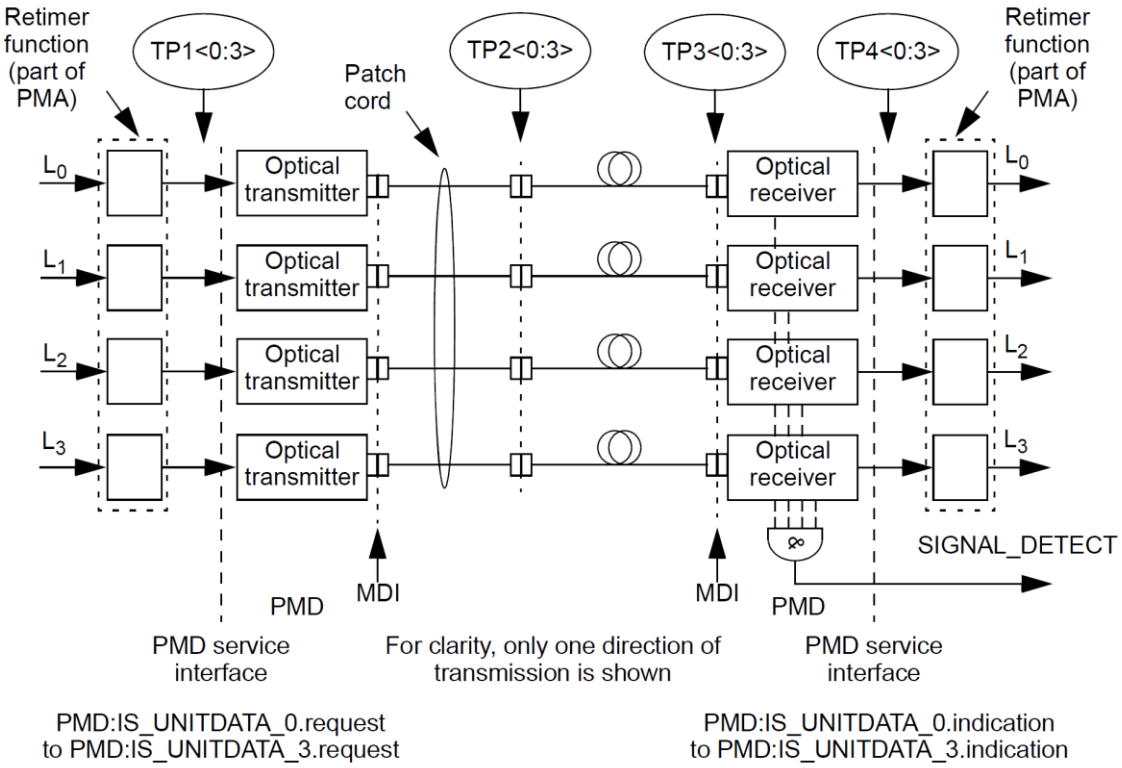


Figure 3-36. *Block diagram for 400GBASE-DR4 transmit/receive paths.*

3.9.1.1 PMD transmit function

The PMD Transmit function shall convert the four symbol streams requested by the PMD service interface messages *PMD:IS_UNITDATA_0.request* to *PMD:IS_UNITDATA_3.request* into four separate optical signals. The four optical signals shall then be delivered to the MDI, which contains four parallel light paths for transmit, according to the transmit optical specifications in this section. The highest optical power level in each signal stream shall correspond to *tx_symbol = three* and the lowest optical power level shall correspond to *tx_symbol = zero*.

3.9.1.2 PMD receive function

The PMD Receive function shall convert the four parallel optical signals received from the MDI into separate symbol streams for delivery to the PMD service interface

using the messages *PMD:IS_UNITDATA_0.indication* to *PMD:IS_UNITDATA_3.indication*, all according to the receive optical specifications in this section. The higher optical power level in each signal stream shall correspond to *rx_symbol = three* and the lowest shall correspond to *rx_symbol = zero*.

3.9.1.3 PMD global signal detect function

The PMD global signal detect function reports the state of *SIGNAL_DETECT* via the PMD service interface. The *SIGNAL_DETECT* parameter is signaled continuously, while the *PMD:IS_SIGNAL.indication* message is generated when a change occurs in the value of *SIGNAL_DETECT*. The *SIGNAL_DETECT* parameter defined in this Section maps to the *SIGNAL_OK* parameter in the inter-sublayer service interface primitives.

SIGNAL_DETECT is a global indicator of the presence of optical signals on all four lanes. The value of the *SIGNAL_DETECT* parameter shall be generated according to the conditions defined in **Table 3-43**. The PMD receiver is not required to verify whether a compliant 400GBASE-DR4 signal is being received. This standard imposes no response time requirements on the generation of the *SIGNAL_DETECT* parameter.

Receive conditions	SIGNAL_DETECT value
For any lane. Average optical power at TP3 ≤ -15dBm	FAIL
For all lanes. [(Optical power at TP3 ≥ average receive power, each lane (min) **Table 3-46**) AND (Compliant 400GBASE–R signal input)]	OK
All other conditions	Unspecified

Table 3-43. *SIGNAL_DETECT value definition.*

All implementations must provide adequate margin between the input optical power level at which the *SIGNAL_DETECT* parameter is set to *OK*, and the inherent noise level of the PMD including the effects of the crosstalk, the power supply noise, etc.

Various implementations of the Signal Detect function are permitted by this standard, including implementations that generate the *SIGNAL_DETECT* parameter values in response to the amplitude of the modulation of the optical signal and implementations that respond to the average optical power of the modulated optical signal.

3.9.1.4 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard. When the MDIO is implemented, each *PMD_signal_detect_i*, where *i* represents the lane number in the range 0:3, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of **Table 3-43**.

3.9.2 Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for 400GBASE-DR4. While it is expected that a PMD will map electrical lane *i* to optical lane *i* and vice versa, there is no need to define the physical ordering of the lanes, as the PCS sublayer is capable of receiving the lanes in any arrangement. The positioning of transmit and receive lanes at the MDI is specified in Section **3.9.7.2**.

3.9.3 PMD to MDI optical specifications for 400GBASE-DR4

The operating range for the 400GBASE-DR4 PMD is defined in **Table 3-44**. A 400GBASE-DR4 compliant PMD operates on type B1.1, B1.3, or B6_a single-mode fibers according to the specifications defined in **Table 3-51**. A PMD that exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., a 400GBASE-DR4 PMD operating at 600m meets the operating range requirement of 2m to 500m).

PMD type	Required operating range
400GBASE-DR4	2m to 500m

Table 3-44. *400GBASE-DR4 operating range.*

3.9.3.1 400GBASE-DR4 transmitter optical specifications

The 400GBASE-DR4 transmitter shall meet the specifications defined in **Table 3-45** per the definitions in Section **3.9.4**.

Description	Value	Unit
Signaling rate, each lane (range)	53.125 ± 100 ppm	GBd
Modulation format	PAM4	-
Lane wavelength (range)	1304.5 to 1317.5	nm
Side-mode suppression ratio (SMSR), (min)	30	dB
Average launch power, each lane (max)	4	dBm
Average launch power, each lane [a] (min)	-2.9	dBm

Outer Optical Modulation Amplitude (OMA_{outer}), each lane (max)	4.2	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (min) [b]	-0.8	dBm
Launch power in OMA_{outer} minus TDECQ, each lane (min)	-2.2	dBm
Transmitter and dispersion eye closure for PAM4 (TDECQ), each lane (max)	3.4	dB
Average launch power of OFF transmitter, each lane (max)	-15	dBm
Extinction ratio, each lane (min)	3.5	dB
$RIN_{21.4}OMA$ (max)	-136	dB/Hz
Optical return loss tolerance (max)	21.4	dB
Transmitter reflectance [c] (max)	-26	dB
[a] Average launch power, each lane (min) is informative and not the principal indicator of signal strength. [b] Even if the TDECQ < 1.4 dB, the OMA_{outer} (min) must exceed these values. [c] Transmitter reflectance is defined looking into the transmitter.		

Table 3-45. *400GBASE-DR4 transmit characteristics.*

3.9.3.2 400GBASE-DR4 receive optical specifications

The 400GBASE-DR4 receiver shall meet the specifications defined in **Table 3-46** per the definitions in Section **3.9.4**.

Description	Value	Unit
Signaling rate, each lane (range)	53.125 ± 100 ppm	GBd
Modulation format	PAM4	-
Lane wavelengths (range)	1304.5 to 1317.5	nm
Damage threshold [a], each lane	5	dBm
Average receive power, each lane (max)	4	dBm
Average receive power, each lane [b] (min)	-5.9	dBm
Receive power (OMA_{outer}), each lane (max)	4.2	dBm
Receiver reflectance (max)	-26	dB
Receiver sensitivity (OMA_{outer}), each lane [c] (max)	-4.4	dBm
Stressed receiver sensitivity (OMA_{outer}), each lane [d] (max)	-1.9	dBm
Conditions of stressed receiver sensitivity test: [e]		
Stressed eye closure for PAM4 (SECQ), lane under test	3.4	dB
OMA_{outer} of each aggressor lane	4.2	dBm

[a] The receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level. The receiver does not have to operate correctly at this input power. [b] Average receive power, each lane (min) is informative and not the principal indicator of signal strength. [c] Receiver sensitivity (OMA_{outer}), each lane (max) is informative and is defined for a transmitter with SECQ of 0.9dB. [d] Measured with conformance test signal at TP3 (see Section **3.9.4.9**) for the BER specified in Section **3.9**. [e] These test conditions are for measuring stressed receiver sensitivity. They are not characteristics of the receiver.

Table 3-46. *400GBASE-DR4 receive characteristics.*

3.9.3.3 400GBASE-DR4 illustrative link power budget

An illustrative power budget and penalties for 400GBASE-DR4 channels are shown in **Table 3-47**.

Parameter	Value	Unit
Power budget (for max TDECQ)	6.5	dB
Operating distance	500	m
Channel insertion loss [a]	3	dB
Maximum discrete reflectance	See Section **3.9.7**	dB
Allocation for penalties [b] (for max TDECQ)	3.5	dB
Additional insertion loss allowed	0	dB
[a] The channel insertion loss is calculated using the maximum distance specified in **Table 3-44** and cabled optical fiber attenuation of 0.5dB/km at 1304.5nm plus an allocation for connection and splice loss given in Section **3.9.7**. [b] Link penalties are used for link budget calculations. They are not requirements and are not meant to be tested.		

Table 3-47. *400GBASE-DR4 illustrative link power budget.*

3.9.4 Definition of optical parameters and measurements methods

All transmitter optical measurements shall be made through a short patch cable, between 2m and 5m in length, unless otherwise specified.

3.9.4.1 Test patterns for optical parameters

While compliance is to be achieved in normal operation, specific test patterns are determined for measurement consistency and the measurement of some parameters.

Table 3-49 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the sections in which each parameter is defined. Any of the test patterns given for a particular test in **Table 3-49** may be used to perform that test. The test patterns used in this Section are shown in **Table 3-48**.

Pattern	Pattern description
Square wave	Square wave (8 threes, 8 zeros)
3	PRBS31Q
4	PRBS31Q
5	Scrambled idle
6	SSPRQ

Table 3-48. *Test patterns.*

Parameter	Pattern	Related section
Wavelength	Square wave, 3, 4, 5, 6 or valid 400GBASE-R signal	**3.9.4.2**
Side mode suppression ratio	3, 5, 6 or valid 400GBASE-R signal	**-**
Average optical power	3, 5, 6 or valid 400GBASE-R signal	**3.9.4.3**
Outer Optical Modulation Amplitude (OMA_{outer})	4 or 6	**3.9.4.4**
Transmitter and dispersion eye closure for PAM4 (TDECQ)	6	**3.9.4.5**
Extinction ratio	4 or 6	**3.9.4.6**
$RIN_{21.4}OMA$	Square wave	**3.9.4.7**
Stressed receiver conformance test signal calibration	6	**3.9.4.9**
Stressed receiver sensitivity	3 or 5	**3.9.4.9**

Table 3-49. *Test-pattern definitions and related sections.*

3.9.4.2 Wavelength

The wavelength of each optical lane shall be within the range given in **Table 3-45** if measured per IEC 61280-1-3. The lane under test is modulated using the test pattern defined in **Table 3-49**.

3.9.4.3 Average optical power

The average optical power of each lane shall be within the limits given in **Table 3-45** if measured using the methods given in IEC 61280-1-1. The average optical power is measured using the test pattern defined in **Table 3-49**, per the test setup in **Figure 3-17**.

3.9.4.4 Outer Optical Modulation Amplitude (OMA_{outer})

The OMA_{outer} of each lane shall be within the limits given in **Table 3-45**. The OMA_{outer} is measured using a test pattern specified for OMA_{outer} in **Table 3-49** as the difference between the average optical launch power level P_3, measured over the central 2 UI of a run of 7 threes, and the average optical launch power level P_0, measured over the central 2 UI of a run of 6 zeros, as shown in **Figure 3-37**.

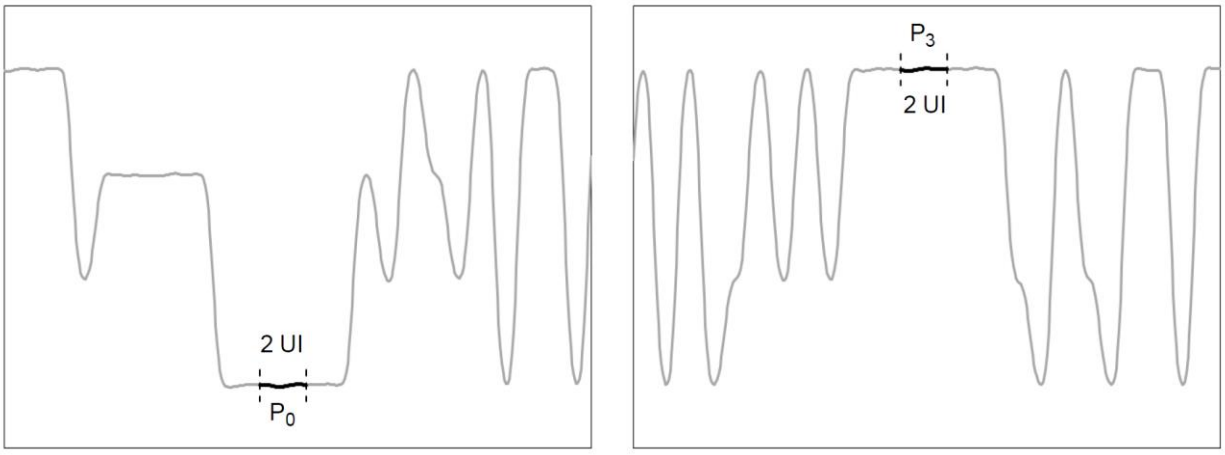


Figure 3-37. *Example power levels P_0 and P_3 from PRBS13Q test pattern.*

3.9.4.5 Transmitter and dispersion eye closure for PAM4 (TDECQ)

The TDECQ of each lane shall be within the limits given in **Table 3-45** if measured using the methods specified in Sections **3.6.4.6**, **3.6.4.7**, and **3.6.4.8** using a reference equalizer as described in Section **3.6.4.9**, with the following exceptions:

- The signaling rate of the test pattern generator is as given in **Table 3-45** and uses the test pattern specified for TDECQ in **Table 3-49**.

- The combination of the O/E converter and the oscilloscope has a fourth-order Bessel-Thomson filter response with a bandwidth of approximately 26.5625GHz.
- The normalized noise power density spectrum *N(f)* is equivalent to white noise filtered by a fourth-order Bessel-Thomson response filter with a bandwidth of 26.5625GHz.

3.9.4.6 Extinction ratio

The extinction ratio of each lane shall be within the limits given in **Table 3-45** if measured using a test pattern specified for extinction ratio in **Table 3-49**. The extinction ratio of a PAM4 optical signal is defined as the ratio of the average optical launch power level P_3, measured over the central 2 UI of a run of 7 threes, and the average optical launch power level P_0, measured over the central 2 UI of a run of 6 zeros, as shown in **Figure 3-37**.

3.9.4.7 Relative intensity noise ($RIN_{21.4}OMA$)

RIN shall be as defined by the measurement methodology of PMD type 10GBASE-S (Short wavelength serial), 10GBASE-L (Long wavelength serial), and 10GBASE-E (Extra-long wavelength serial) with the following exceptions:

a. The optical return loss is 21.4dB.
b. Each lane may be tested individually with the sum of the optical power from all of the lanes not under test being below –30dBm.
c. The upper –3dB limit of the measurement apparatus is to be approximately equal to the signaling rate (i.e., 53.2GHz).
d. The test pattern is according to **Table 3-49**.

3.9.4.8 Receiver sensitivity

Receiver sensitivity, which is defined for an input signal with SECQ of 0.9dB (e.g., an ideal input signal without overshoot), is informative and compliance is not required. If measured, the test signal should have negligible impairments such as ISI, rise/fall times, jitter and RIN. Instead, the normative requirement for receivers is stressed receiver sensitivity.

3.9.4.9 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in **Table 3-46** if measured using the method defined in Section **3.6.4.13** with the following exceptions:

- The SECQ of the stressed receiver conformance test signal is measured according to Section **3.9.4.5**, except that the test fiber is not used.
- The signaling rate of the test pattern generator and the extinction ratio of the E/O converter are as given in **Table 3-45** using test patterns specified in **Table 3-49**.
- The required values of the “Stressed receiver sensitivity (OMA_{outer}), each lane (max)”, “Stressed eye closure for PAM4 (SECQ), lane under test”, and “OMA_{outer} of each aggressor lane” are as given in **Table 3-46**.

3.9.5 Safety, installation, environment, and labeling

All equipment subject to this Section shall conform to IEC 60950-1.

400GBASE-DR4 optical transceivers shall conform to Hazard Level 1 laser requirements as defined in IEC 60825-1 and IEC 60825-2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product’s laser, safety features, labeling, use, maintenance, and service.

It is recommended that proper installation practices, as defined by applicable local codes and regulation, be followed in every instance in which such practices are applicable.

Normative specifications in this Section shall be met by a system integrating a 400GBASE-DR4 PMD over the life of the product while the product operates within the manufacturer’s range of environmental, power, and other specifications.

It is recommended that manufacturers indicate in the literature associated with the PHY the operating environmental conditions to facilitate selection, installation, and maintenance. It is recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of this Section will be met.

A system integrating a 400GBASE-DR4 PMD shall comply with applicable local and national codes for the limitation of electromagnetic interference.

3.9.6 Fiber optic cabling model

The fiber optic cabling model is shown in **Figure 3-38**.

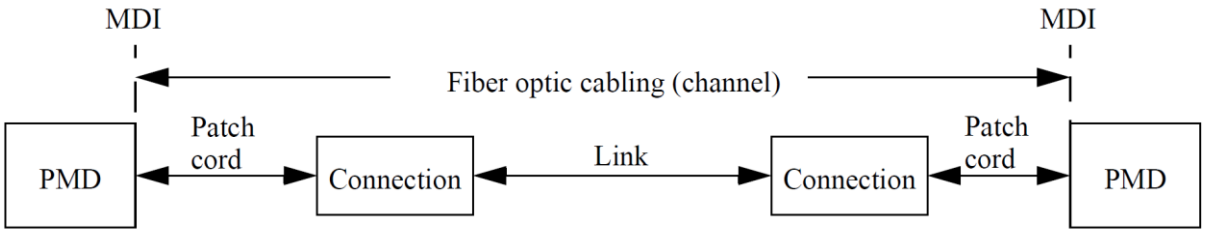


Figure 3-38. *Fiber optic cabling model.*

The channel insertion loss is given in **Table 3-50**. A channel contains additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion comply with the specifications. Insertion loss measurements of installed fiber cables are made in compliance with IEC 61280-4-1. The fiber optic cabling model (channel) determined here is the same as a simplex fiber optic permanent link. The term channel is used here for consistency with generic cabling standards.

Description	400GBASE-DR4	Unit
Operating distance (max)	500	m
Channel insertion loss [a, b] (max)	3	dB
Channel insertion loss (min)	0	dB
Positive dispersion [b] (max)	0.8	ps/nm
Negative dispersion [b] (min)	-0.93	ps/nm
DGD_max [c]	2.24	ps
Optical return loss (min)	37	dB
[a] These channel insertion loss values include cable, connectors, and splices. [b] Over the wavelength range 1304.5nm to 1317.5nm. [c] Differential Group Delay (DGD) is the time difference at reception between the fractions of a pulse that were transmitted in the two principal states of polarization of an optical signal, where *DGD_max* is the maximum differential group delay that the system must tolerate.		

Table 3-50. *Fiber optic cabling (channel) characteristics.*

3.9.7 Characteristics of the fiber optic cabling (channel)

The 400GBASE-DR4 fiber optic cabling shall meet the specifications defined in **Table 3-51**. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

The fiber optic cable requirements are satisfied by cables containing IEC 60793-2-50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-

mode), or type B6_a (bend insensitive) fibers or the requirements in **Table 3-51** where they differ.

Description	Value	Unit
Nominal fiber specification wavelength	1310	nm
Cabled optical fiber attenuation (max)	0.5 [a]	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0.093	ps/nm^{2}km
[a] The 0.5dB/km attenuation is provided for Outside Plant cable as defined in CENELEC EN 50173-1 and ISO/IEC 11801-1.		

Table 3-51. *Optical fiber and cable characteristics.*

An optical fiber connection, as shown in **Figure 3-38**, consists of a mated pair of optical connectors.

The maximum link distance is based on an allocation of 2.75dB total connection and splice loss. For example, this allocation supports five connections with an average insertion loss per connection of 0.5dB. Connections with different loss characteristics may be used provided the requirements of **Table 3-50** are met.

The maximum value for each discrete reflectance shall be less than or equal to the value shown in **Table 3-52** corresponding to the number of discrete reflectances above –55dB within the channel. For numbers of discrete reflectances in between two numbers shown in the table, the lower of the two corresponding maximum discrete reflectance values applies.

Number of discrete reflectances above –55 dB	Maximum value for each discrete reflectance
1	-37dB
2	-42dB
4	-45dB
6	-47dB
8	-48dB
10	-49dB

Table 3-52. *Maximum value of each discrete reflectance.*

3.9.7.1 Medium Dependent Interface (MDI)

The 400GBASE-DR4 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown **Figure 3-38**). The 400GBASE-DR4 PMD is coupled to the fiber optic cabling through one connector

plug into the MDI optical connector as shown in **Figure 3-39**. Example constructions of the MDI include the following:

a. PMD with a connectorized fiber pigtail plugged into an adapter.
b. PMD connector.

3.9.7.2 Optical lane assignments

The four transmit and four receive optical lanes of 400GBASE-DR4 shall occupy the positions depicted in **Figure 3-39** when looking into the MDI connector with the keyway feature on top. The interface contains eight active lanes within twelve total positions. The transmit optical lanes occupy the left-most four positions and the receive optical lanes occupy the right-most four positions. The four center positions are unused.

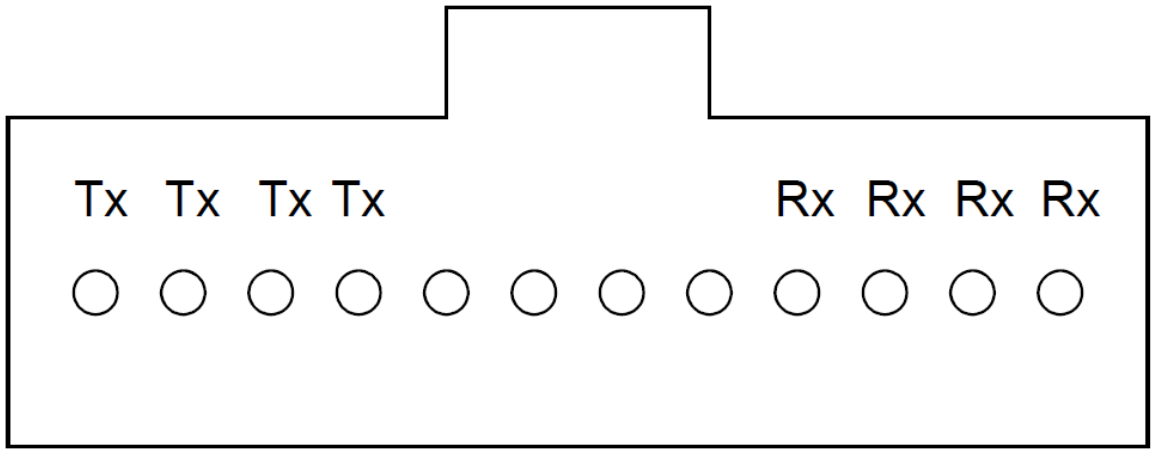


Figure 3-39. *400GBASE-DR4 optical lane assignments.*

3.9.7.3 MDI requirements

The MDI shall meet the dimensional specifications of IEC 61754-7-1 interface 7-1-9: MPO device connector, angled interface. The plug terminating the optical fiber cabling shall meet the dimensional specifications of IEC 61754-7-1 interface 7-1-1: MPO female plug connector, down-angled interface for 2 to 12 fibres. The MDI shall optically mate with the plug on the optical fiber cabling. **Figure 3-40** shows an MPO female plug connector with down-angled interface, and an MDI as an active device connector with angled interface.

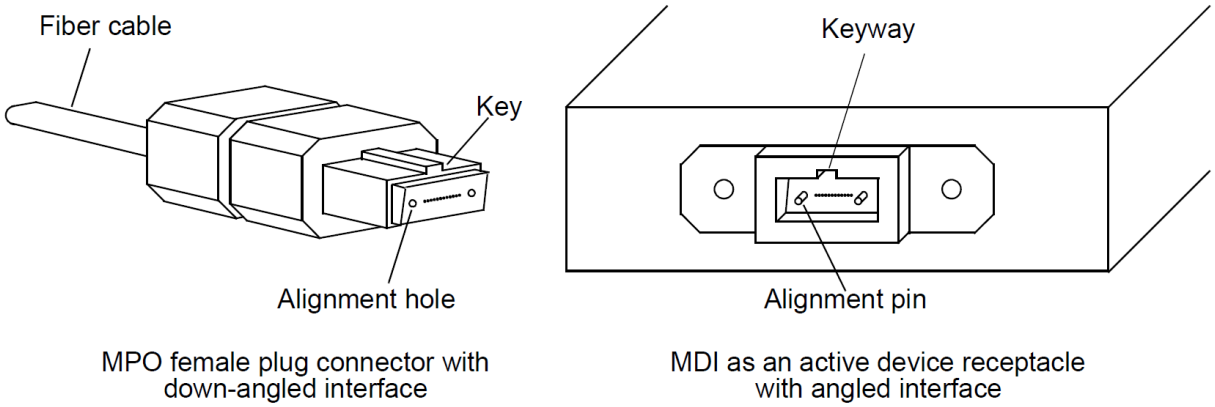


Figure 3-40. *MPO female plug with down-angled interface and MDI active device connector with angled interface.*

The MDI shall meet the interface performance specifications of IEC 61753-021-2 for performance level D/2.

NOTE: Transmitter compliance testing is performed at TP2 as defined in Section **3.8.1**, not at the MDI.

4 Architectural Perspectives of 800Gb/s and 1.6Tb/s Ethernet

The IEEE 802.3 Beyond 400Gb/s Ethernet Task Force (B400G TF) considers the need of a broad range of PHY specifications and simultaneously the development of the architecture to support the next generation of Ethernet applications, in both MAC rates and lane signaling rates. This will enable a “building block” technologies approach to address multiple lanes and media for the industry that has a growing acceptance of multi-lane infrastructures. (D'Ambrosia, Draft Project Documentation Update, 2021)

The project of developing the 800Gb/s and 1.6Tb/s Ethernet transmission was approved initially as amendment **IEEE P802.3df, Media Access Control Parameters, Physical Layers and Management Parameters for 200Gb/s, 400Gb/s, 800Gb/s and 1.6Tb/s Operation** in September 2021. In July 2022, the emerging baseline status indicated that 100Gbps/lane, 200Gbps/lane, and LR/ER objectives were on different timelines, so the B400G TF decided to split the IEEE P802.3df into two projects: (D’Ambrosia & Nowell, Splitting IEEE P802.3df, Proposed Draft IEEE P802.3df PAR Modification Responses 400 GbE and 800 GbE Objectives, 2022)

- 100Gbps/lane-based standard as **IEEE 802.3df, Media Access Control Parameters for 800 Gb/s and Physical Layers and Management Parameters for 400 Gb/s and 800 Gb/s Operation**. The IEEE 802.3df is estimated to be released in June 2024.
- 200Gbps/lane-based standard as **IEEE P802.3dj, Media Access Control Parameters for 1.6 Tb/s and, Physical Layers and Management Parameters for 200 Gb/s, 400 Gb/s, 800 Gb/s, and 1.6 Tb/s Operation**. The IEEE 802.3dj is estimated to be released in March 2026.

The complete standard will define Ethernet local area, access, and metropolitan area networks. Ethernet will be specified at selected speeds of operation, and uses a common MAC specification and MIB. The CSMA/CD MAC protocol will also specify shared medium (half duplex) operation, as well as full duplex operation. The speed specific MIIs will provide an architectural and optional implementation interface to selected PHY entities. The PHY will encode frames for transmission and will decode received frames with the modulation specified for the speed of operation, transmission medium and supported link length. Other specified capabilities will include the control

and the management protocols, as well as the provision of power over selected twisted pair PHY types. (D'Ambrosia & Nowell, Splitting IEEE P802.3df, Proposed Draft IEEE P802.3df PAR Modification Responses 400 GbE and 800 GbE Objectives, 2022)

The IEEE P802.3df will define the Ethernet MAC parameters, the PHY specifications, and the management parameters for the transfer of the Ethernet format frames at 800Gb/s over copper, multi-mode fiber, and single-mode fiber based on 100 Gb/s signaling technology, and use this work to define derivative physical layer specifications and management parameters for the transfer of Ethernet format frames at 400 Gb/s. (D'Ambrosia & Nowell, Splitting IEEE P802.3df, Proposed Draft IEEE P802.3df PAR Modification Responses 400 GbE and 800 GbE Objectives, 2022)

The IEEE P802.3dj will define the Ethernet MAC parameters for 1.6 Tb/s. The IEEE P802.3dj will also define the PHY specifications, and management parameters for the transfer of Ethernet format frames at 800Gb/s and 1.6Tb/s over copper, multi-mode fiber, and single-mode fiber PMDs based on greater than 100 Gb/s signaling technologies. This work will be used to define derivative PHY specifications and management parameters for the transfer of Ethernet format frames at 200 Gb/s and 400 Gb/s. (D'Ambrosia & Nowell, Splitting IEEE P802.3df, Proposed Draft IEEE P802.3dj, PAR Responses200 GbE, 400 GbE, 800 GbE and 1.6 TbE Objectives, 2022)

The adopted objectives at this stage of the proposed IEEE P802.3df are: (D'Ambrosia & Nowell, Proposed Objectives –Modified P802.3df PAR, 2022)

i. For non-rate specific:
    - Support full-duplex operation only.
    - Preserve the Ethernet frame format utilizing the Ethernet MAC.
    - Preserve minimum and maximum FrameSize of current IEEE 802.3 standard.
    - Support a BER of better than or equal to 10^{-13} at the MAC/PLS service interface (or the frame loss ratio equivalent).
    - Provide support to enable mapping over Optical Transport Network (OTN).
ii. For 400Gb/s related:
    - Support a MAC data rate of 400Gb/s.

- Define a PHY specification that supports 400Gb/s operation over 4 pairs of SMF with lengths up to at least 2km.

iii. For 800Gb/s related:

- Support a MAC data rate of 800Gb/s.
- Support optional eight-lane 800Gb/s AUIs for chip-to-module and chip-to-chip applications.
- Define a PHY specification that supports 800Gb/s operation:
  - over eight lanes of twin axial copper cables with a reach up to at least 2m.
  - over eight lanes over electrical backplanes (BPs) supporting an insertion loss ≤ 28dB at 26.56GHz.
  - over 8 pairs of MultiMode Fiber (MMF) with lengths up to at least 50m.
  - over 8 pairs of MMF with lengths up to at least 100m.
  - over 8 pairs of SMF with lengths up to at least 500m.
  - over 8 pairs of SMF with lengths up to at least 2km.

The adopted objectives at this stage of the proposed IEEE P802.3dj are: (D'Ambrosia & Nowell, Proposed Objectives – P802.3dj PAR, 2022)

i. For non-rate specific:

- Support full-duplex operation only.
- Preserve the Ethernet frame format utilizing the Ethernet MAC.
- Preserve minimum and maximum FrameSize of current IEEE 802.3 standard.
- Support a BER of better than or equal to 10^{-13} at the MAC/PLS service interface (or the frame loss ratio equivalent).
- Provide support to enable mapping over OTN.

ii. For 200Gb/s related:

- Support a MAC data rate of 200Gb/s.
- Support optional single-lane 200Gb/s AUIs for chip-to-module and chip-to-chip applications.
- Define a PHY specification that supports 200Gb/s operation:
  - over 1 pair of copper twin-axial cables in each direction with a reach of up to at least 1m.

- over 1 pair of Single-Mode Fiber (SMF) with lengths up to at least 500m.
- over 1 pair of SMF with lengths up to at least 2km.

iii. For 400Gb/s related:

- Support a MAC data rate of 400Gb/s.
- Support optional two-lane 400Gb/s AUIs for chip-to-module and chip-to-chip applications.
- Define a PHY specification that supports 400Gb/s operation:
  - over 2 pairs of copper twin-axial cables in each direction with a reach of up to at least 1m.
  - over 2 pairs of SMF with lengths up to at least 500m.

iv. For 800Gb/s related:

- Support a MAC data rate of 800Gb/s.
- Support optional four-lane 800Gb/s AUIs for chip-to-module and chip-to-chip applications.
- Define a PHY specification that supports 800Gb/s operation:
  - over 4 pairs of copper twin-axial cables in each direction with a reach of up to at least 1m.
  - over 4 pairs of SMF with lengths up to at least 500m.
  - over 4 pairs of SMF with lengths up to at least 2km.
  - over 4 wavelengths over a single SMF in each direction with lengths up to at least 2km.
  - over a single SMF in each direction with lengths up to at least 10km.
  - over a single SMF in each direction with lengths up to at least 40km.

v. For 1.6Tb/s related:

- Support a MAC data rate of 1.6Tb/s.
- Support optional sixteen-lane 1.6Tb/s AUIs for chip-to-module and chip-to-chip applications.
- Support optional eight-lane 1.6Tb/s AUIs for chip-to-module and chip-to-chip applications.
- Define a PHY specification that supports 1.6Tb/s operation:

- over 8 pairs of copper twin-axial cables in each direction with a reach of up to at least 1m.
- over 8 pairs of SMF with lengths up to at least 500m.
- over 8 pairs of SMF with lengths up to at least 2km.

The PHY objectives landscape of the complete standard are described in **Table 4-1**. (D'Ambrosia, Objectives, 2021)

Ethernet Rate	Assumed Signaling Rate	AUI	BackPlane (BP)	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200Gb/s	200Gb/s	Over 1 lane (200GAUI-1)		Over 1 pair (800GBASE-CR1)			Over 1 pair (TBD)	Over 1 pair (TBD)		
400Gb/s	200Gb/s	Over 2 lanes (400GAUI-2		Over 2 pairs (400GBASE-CR2)			Over 2 pairs (TBD)			
800Gb/s	100Gb/s	Over 8 lanes (800GAUI-8)	Over 8 lanes (800GBASE-KR8)	Over 8 pairs (800GBASE-CR8)	Over 8 pairs (800GBASE-VR8)	Over 8 pairs (800GBASE-SR8)	Over 8 pairs (TBD)	Over 8 pairs (TBD)		
	200Gb/s	Over 4 lanes (800GAUI-4)		Over 4 pairs (800GBASE-CR4)			Over 4 pairs (TBD)	1. Over 4 pairs (TBD) 2. over 4 λ's (TBD)		
	TBD								Over single SMF in each direction (TBD)	Over single SMF in each direction (TBD)
1.6Tb/s	100Gb/s	Over 16 lanes (1.6TAUI-16)								
	200Gb/s	Over 8 lanes (1.6TAUI-8)		Over 8 pairs (1.6TBASE-CR8)			Over 8 pairs (TBD)	Over 8 pairs (TBD)		

Table 4-1. *Adopted physical layer objectives landscape.*

TBD means "to be defined". The signaling technologies across Ethernet rates which leveraged by B400G TF are described in **Figure 4-1**:

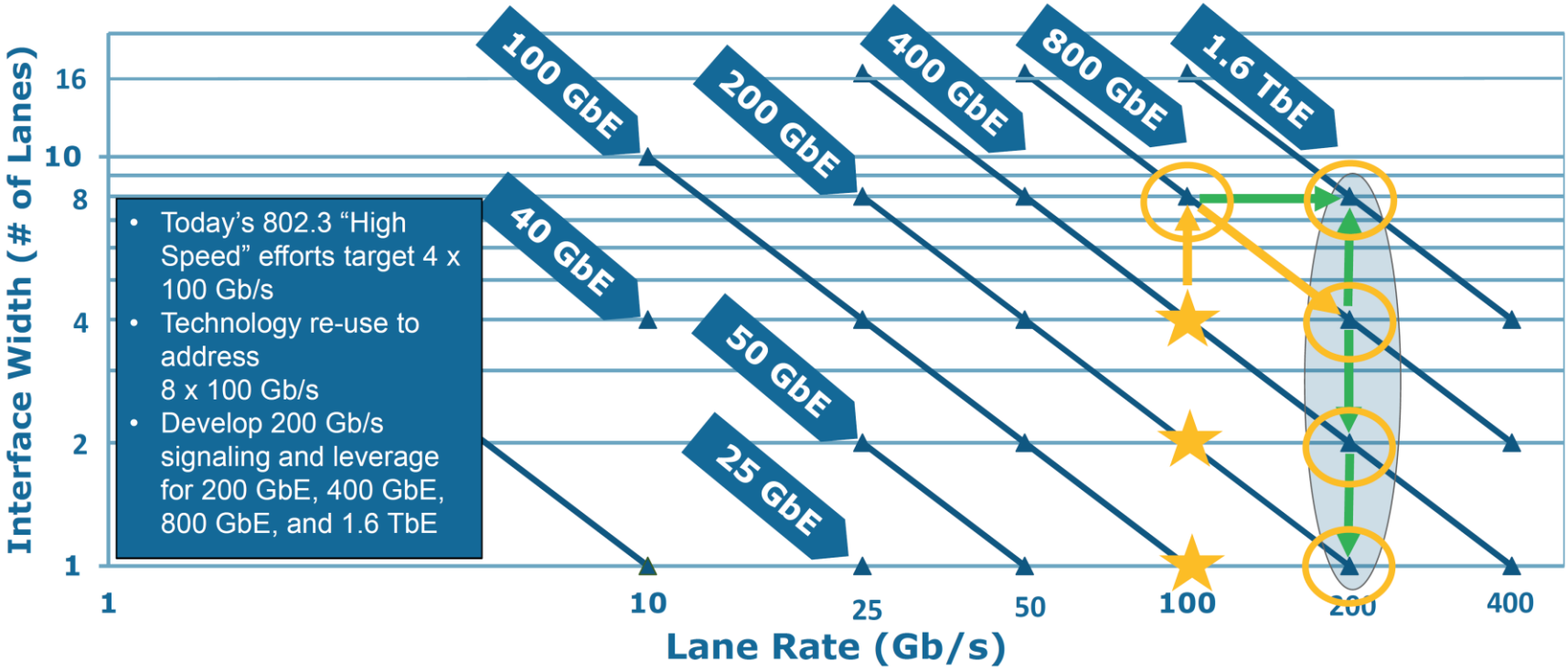
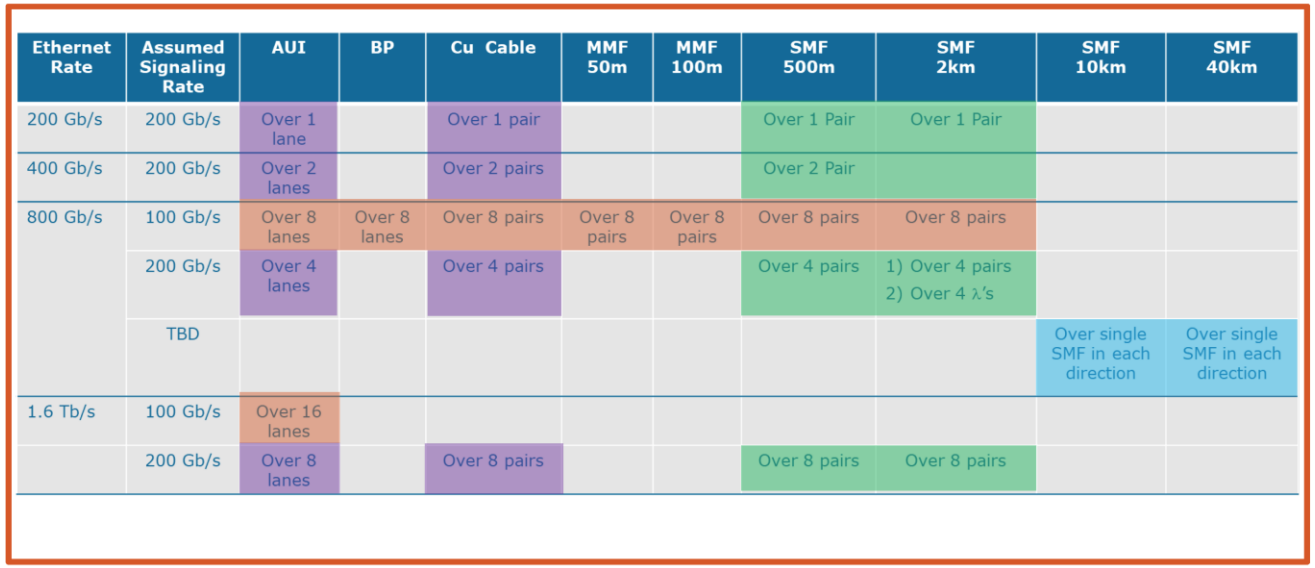


Figure 4-1. *Leveraged Signaling Technologies Across Ethernet Rates.*

The IEEE 802.3df architecture will be based on AUI's using 100Gbps/lane (see the asterisks in **Figure 4-1**), so 400Gb/s will be implemented by using 4 lanes of 100Gb/s and 800Gb/s will be implemented by using 8 lanes of 100Gb/s. Also, the IEEE 802.3dj architecture will be based on AUI's using 200Gbps/lane, so 800Gb/s will be implementing by using 4 lanes of 200Gb/s and 1.6Tb/s will be implementing by using 8 lanes of 200Gb/s. The complete standard needs to be flexible to address all PHY specifications by leveraging technology reuse. **Figure 4-2** illustrates the adopted PHY objectives and the need of the technology reuse.



Ethernet Rate	Assumed Signaling Rate	AUI	BP	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200 Gb/s	200 Gb/s	Over 1 lane		Over 1 pair			Over 1 Pair	Over 1 Pair		
400 Gb/s	200 Gb/s	Over 2 lanes		Over 2 pairs			Over 2 Pair			
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 lanes	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs		
	200 Gb/s	Over 4 lanes		Over 4 pairs			Over 4 pairs	1) Over 4 pairs 2) Over 4 λ's		
	TBD								Over single SMF in each direction	Over single SMF in each direction
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes		Over 8 pairs			Over 8 pairs	Over 8 pairs		

Figure 4-2. *Adopted Physical Layer Objectives and technology reuse.*

Additional objectives to be considered are:

- power,
- backwards/forward interoperability,
- FEC Coding Gain needed per PHY specifications, and
- latency.

4.1 800Gb/s and 1.6TGb/s MAC sublayer

The MAC sublayer, which responds to Layer 2 of the OSI model, is connected to the media (copper or optical) by an Ethernet PHY device, which corresponds to Layer 1 of the OSI model. The PHY device of 100Gbp/s (see **Figure 2-4**), 200Gb/s and 400Gb/s (see **Figure 3-1**) consist of a PMD sublayer, a PMA sublayer, and a PCS sublayer.

For 800Gb/s the MAC sublayer will remain similar as 100Gb/s (IEEE P802.3ck & IEEE 802.3cu™) and 200/400Gb/s (IEEE 802.3bs™). For 1.6TGb/s the MAC sublayer will remain similar as 200/400Gb/s (IEEE 802.3bs™). (Wang, He, & Ren, Technical Feasibility of Logic Layer to Support Rate Objective, 2021)

4.2 Reconciliation Sublayer (RS) and Media Independent Unit (MII)

The RS/MII sublayers will remain similar as Section **2.3** and Section **3.2** with 64-bit data and 8-bit control, which seems to be reasonable. The RS adapts the bit serial protocols of the MAC sublayer to the parallel format of the PCS service interface. This serial to parallel adaptation has a mapping purpose, where the transmit and receive signals are reconciled at the MII to the MAC sublayer. Each direction (transmit (TXD), receive (RXD)) uses 64-bit data signals (TXD<63:0> and RXD<63:0>), 8 control signals (TXC<7:0> and RXC<7:0>), and a clock (TX_CLK and RX_CLK) as shown in **Figure 4-3** and **Figure 3-7**. (IEEE Std 802.3™, 2018)

200GMII is the transparent signal interface between the RS and the PCS for 200Gb/s operation, while 400GMII is the transparent signal interface between the RS and the PCS for 400Gb/s operation.

Further extending RS and MII data bus to larger than 64-bit, e.g., 124-bit/16 Byte is not doable, because this may violate the Deficit Idle Counter (DIC) mechanism and the minimum InterPacket Gap (IPG) requirement of 12 Byte, compromising the line rate

transmission in Ethernet. (Wang, He, & Ren, Technical Feasibility of Logic Layer to Support Rate Objective, 2021)

Notes:

a. For 7nm node ASIC, the 640-bit at 1.29GHz parallel implementation is achievable to enable 800Gb/s 64B/66B encode/decode.

b. Forecasting 7nm and ≤5nm node ASIC, the 1280-bit at 1.29GHz parallel implementation is feasible to enable 1.6Tb/s 64B/66B encode/decode.

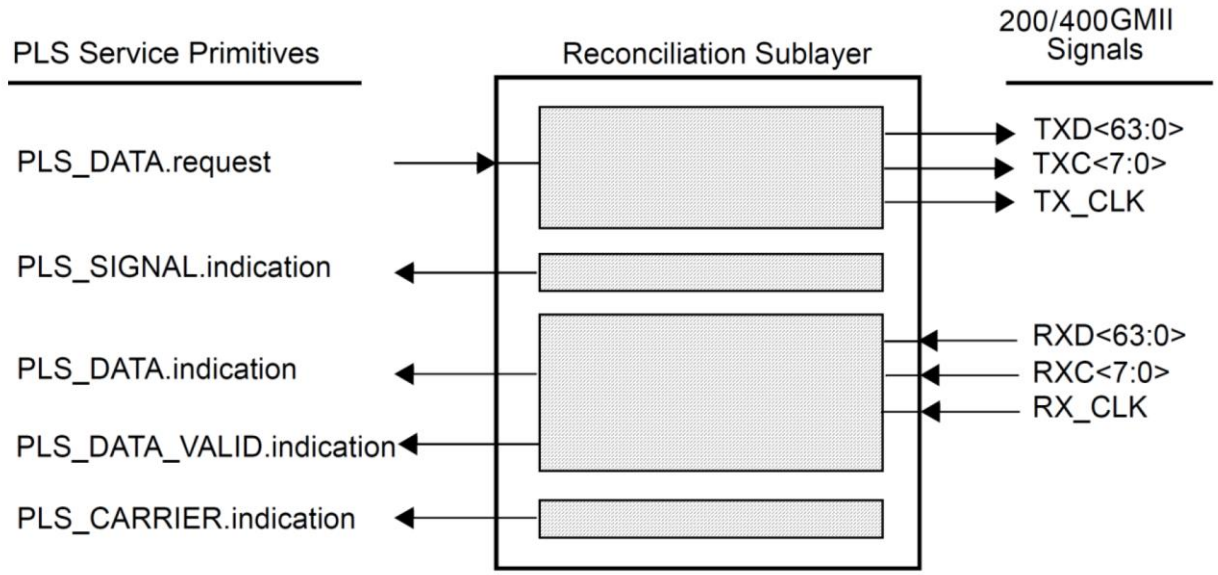


Figure 4-3. *Reconciliation sublayer for 200 and 400Gb/s Ethernet.*

4.3 Physical Coding Sublayer (PCS) and Forward Error Correction (FEC)

The PCS is a sublayer used in certain port types to couple the MII and the PMA. The PCS is responsible for the encoding of data bits into code groups for transmission via the PMA and the subsequent decoding of these code groups from the PMA. IEEE also developed a low-overhead multilane distribution scheme for the PCS of 100Gb/s, 200Gb/s and 400Gb/s Ethernet. In general, 100Gb/s, 200GBASE-R and 400GBASE-R PCSs provide the services required by the CGMII, 200GMII and 400GMII, including the following: (IEEE Std 802.3™, 2018)

a. Encoding/decoding of eight CGMII/200GMII/400GMII data octets to/from 66-bit blocks (64B/66B).

b. Transcoding from 66-bit blocks to/from 257-bit blocks.
c. Reed-Solomon encoding/decoding the 257-bit blocks.
d. Transferring encoded data to/from the PMA.

The PCS has been designed to support all PHY types for 100Gb/s, 200Gb/s and 400Gb/s Ethernet and it is flexible and scalable to support any future PHY types that may be developed, based on future advances in optical transmission. The PCS layer also performs the following functions:

a. Delineation of frames.
b. Transport of control signals.
c. Ensures necessary clock transition density needed by the physical optical and electrical technology.
d. Stripes and re-assembles the information across multiple lanes.

The PCS leverages the 64B/66B coding scheme that was used in 10Gb/s Ethernet and provides a number of useful properties, including low overhead and sufficient code space to support necessary code words.

The multilane distribution scheme developed for the PCS is fundamentally based on a striping of the 66-bit blocks across multiple lanes.

The transmit PCS, therefore, performs the initial 64B/66B encoding and scrambling on the aggregate channel (100Gb/s) before distributing 66-bit block in a round robin basis across the multiple lanes, referred to as “PCS lanes,” as shown in **Figure 4-4**. (D’Ambrosia, Law, & Nowell, 40 Gigabit Ethernet and 100 Gigabit Ethernet. Technology Overview, 2010)

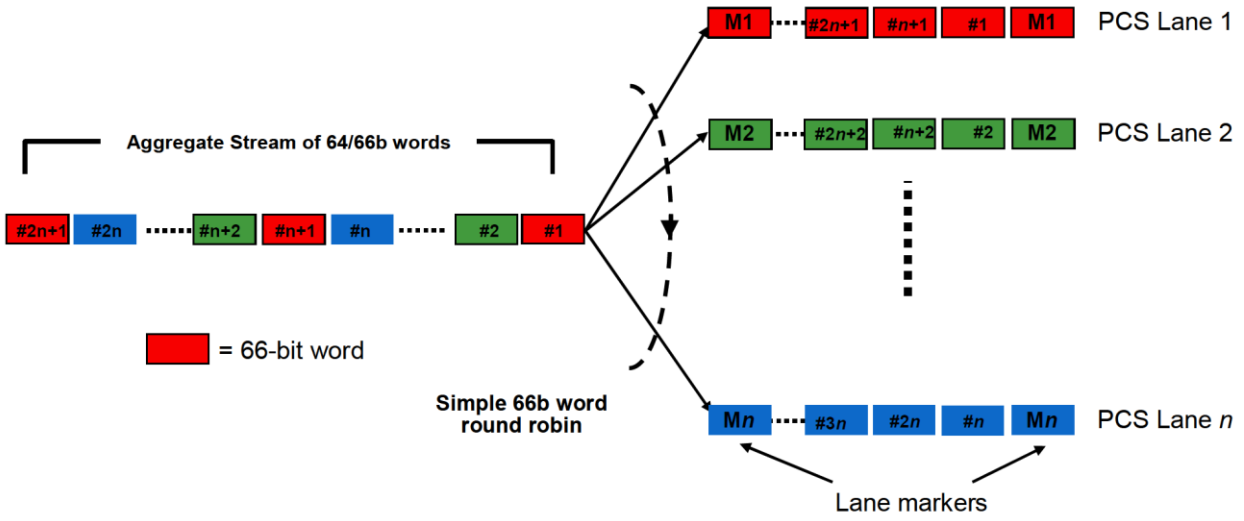


Figure 4-4. *PCS Multilane Distribution Concept.*

In addition, the transmit PCS is transcoded from 64B/66B to 256B/257B encoding and scrambling on the aggregate channel (200Gb/s or 400Gb/s), reducing the overhead and making room for the FEC. The 256B/257B encoded data is then FEC encoded before being transmitted. Data distribution is introduced to support multiple lanes in the PHY. Part of the distribution includes the periodic insertion of an alignment marker, which allows the receive PCS to align data from multiple lanes (see Section **3.4**).

An example of MAC/PCS/PMA using 8 x 100Gb/s Serializer/Deserializer (SerDes) is shown in ***Figure 4-5***: (Shrikhande, 2021)

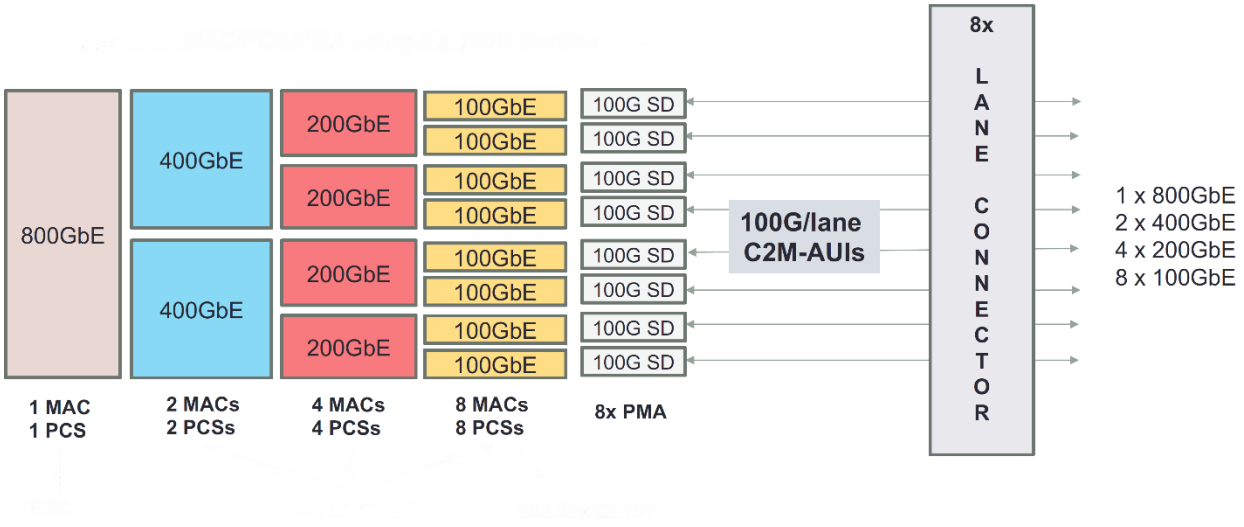


Figure 4-5. *Current 100Gb/s per lane AUI generation.*

IEEE 802.3ck is specifying 100/200/400Gb/s Ethernet using 100Gb/s per lane at the same time.

In addition, an example of MAC/PCS/PMA using 8 x 200Gb/s SerDes is shown in ***Figure 4-6***: (Shrikhande, 2021)

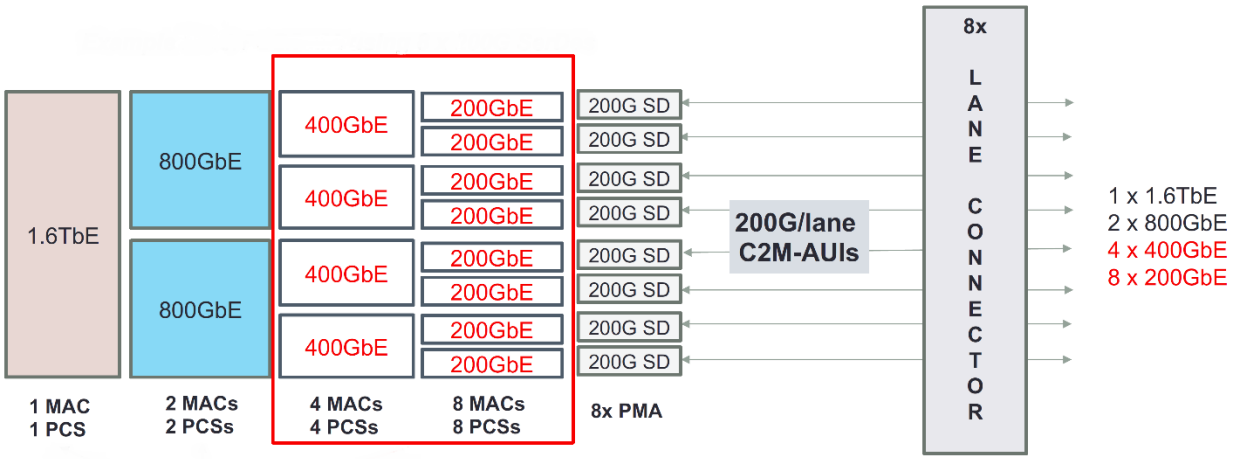


Figure 4-6. *Future 200G/lane AUI generation.*

Ethernet Technology Consortium also proposes that 800Gb/s capability will be supported by utilizing two 400 Gb/s PCSs (with the included FEC) and supporting 32 PCS lanes, each at 25Gb/s. **Figure 4-7** shows at a high level the transmit PCS data flow and functions. The 64B/66B encoder must run at 800Gb/s to create a coherent stream of data, but the rest of the processing is on a 400Gb/s slice of the data. The distribution, as shown in **Figure 4-7**, is based on 1x66b blocks. The alignment marker insertion must be coordinated between the two stacks to ensure that coherent data is received and able to be processed on the receive side. Also, the alignment markers that are inserted will vary when compared to 400Gb/s to enable reception and deskew of a coherent data stream. 2x16 PCS lanes are generated from the two PCS stacks and then they are 4:1 bit multiplexed by the PMA towards the PMD in order to create 8x100Gb/s PMD lanes.

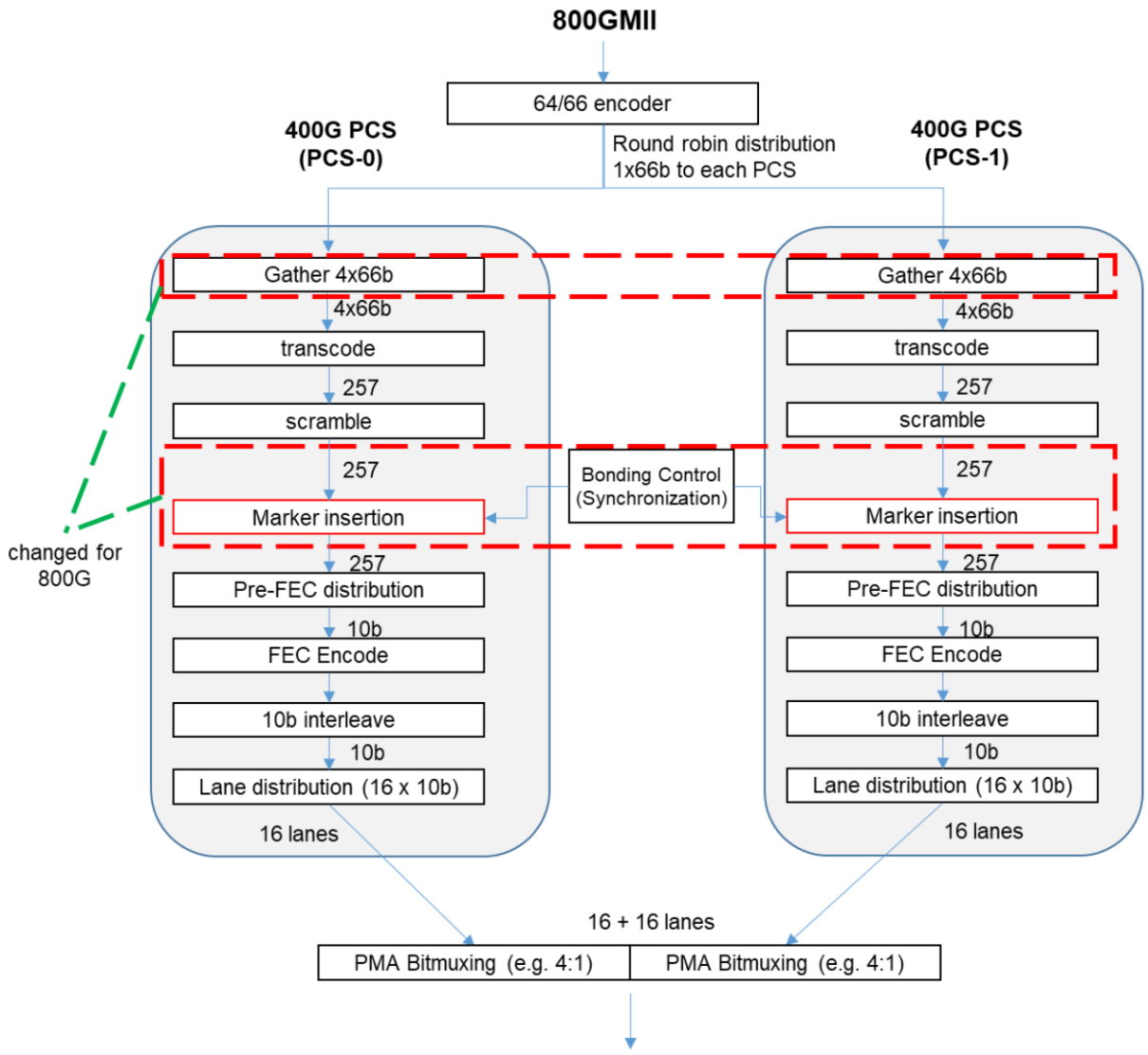


Figure 4-7. *800Gb/s PCS transmit flow.*

64B/66B encoding is performed on the full 800Gb/s data stream as described in Section **3.4**. 66B blocks will be distributed to the two PCS instances 1x66B block at a time, in a round robin fashion, starting with PCS-0 and then to PCS-1 and back to PCS-0 again, as shown in **Figure 4-8**.

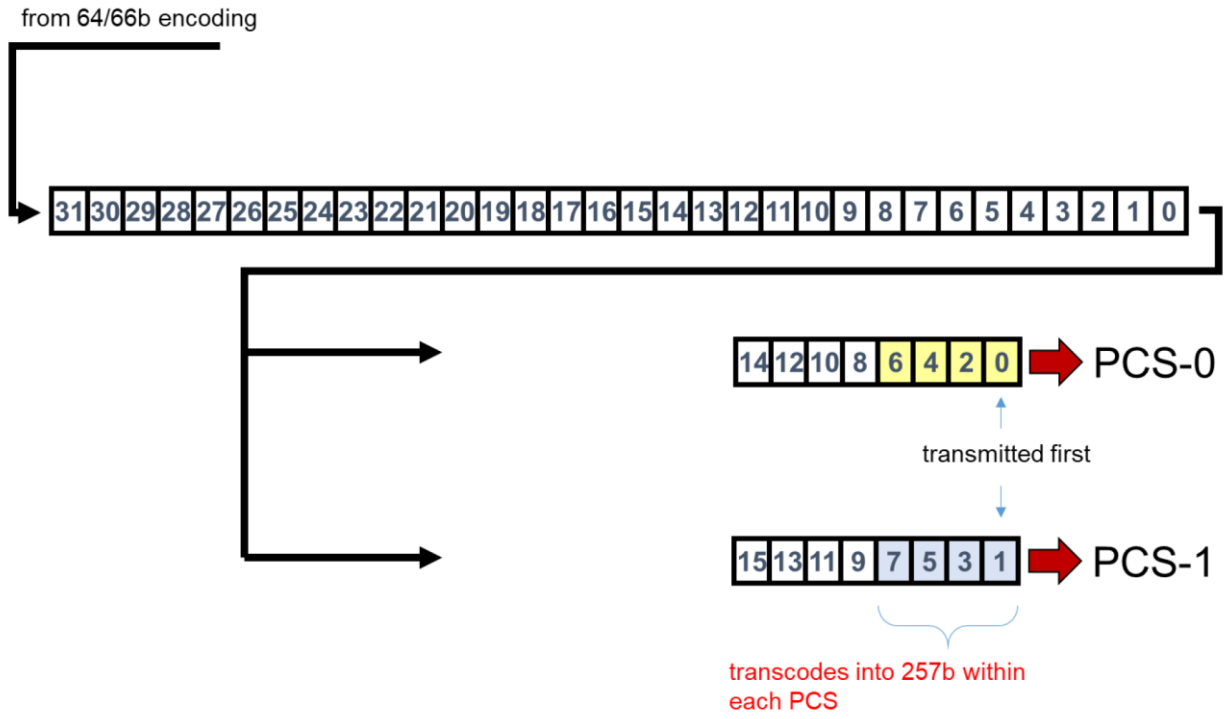


Figure 4-8. *66B block Round Robin distribution.*

The receive PCS flow is shown in **Figure 4-9**. It is just the reverse processing when compared to the transmit PCS, and still focused mostly on 2x400Gb/s protocol stacks. This 800Gb/s definition allows any PCS lane to be received on any PMD lane, and so 32 PCS lanes must be locked to, and reordered appropriately to recover the data.

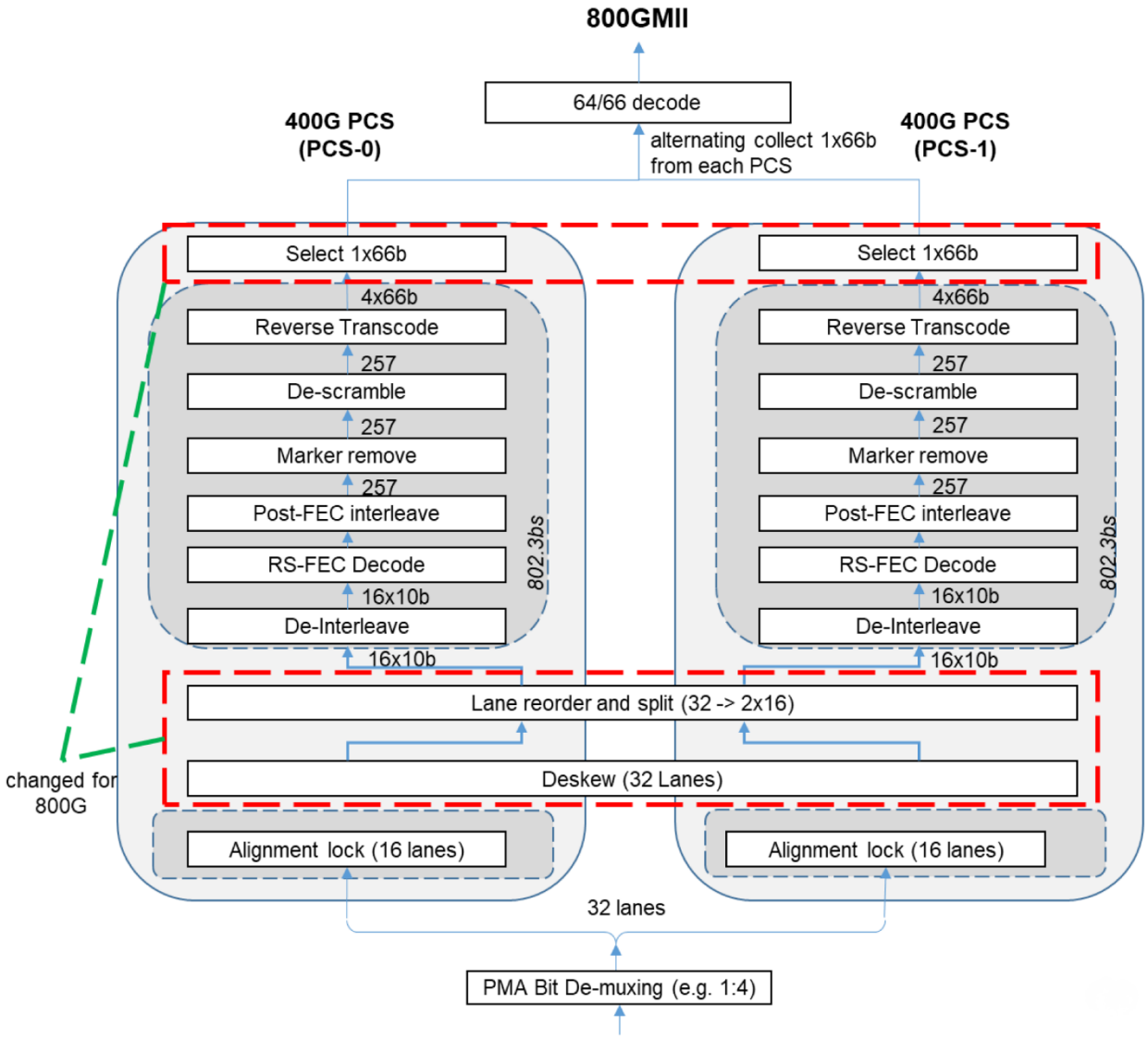


Figure 4-9. *800Gb/s PCS receive flow.*

1.6Tb/s capability will be supported by utilizing 4 x 400 Gb/s PCSs with FEC, which uses Round Robin distribution across the PCS lanes, as described in **Figure 4-10**. (Theodoras, 2021)

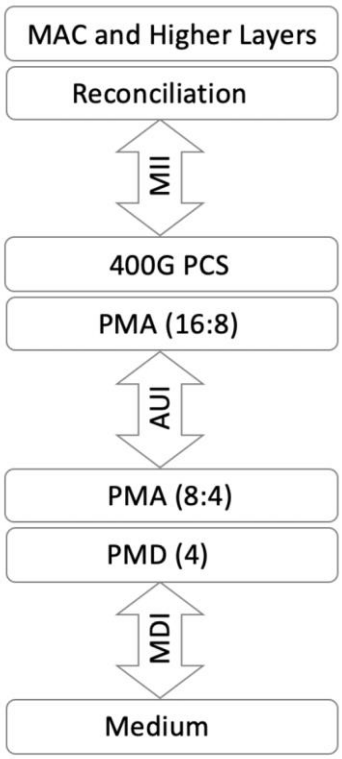


Figure 4-10. *400 Gb/s PCSs.*

The FEC sublayer can be placed in between the PCS and PMA sublayers or between two PMA sublayers. The FEC is initiated and operates autonomously for each PCS lane. Both 200GBASE-R and 400GBASE-R are using the 64B/66B code. The 64B/66B code supports transmission of data and control characters and it transcoded to 256B/257B encoding to reduce the overhead and make room for FEC. The 256B/257B encoded data are then FEC encoded before being transmitted. Data distribution is introduced to support multiple lanes in the PHY. (IEEE Std 802.3™, 2018)

The development of new Ethernet transmission rates at 800Gb/s and 1.6Tb/s needs the adoption of an initial holistic approach which may support one or more FEC schemes. The IEEE 802.3 B400G TF introduces a new architecture based on one or more FEC schemes that may be applied to 800Gb/s and 1.6Tb/s Ethernet rates, as shown in **Figure 4-11**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

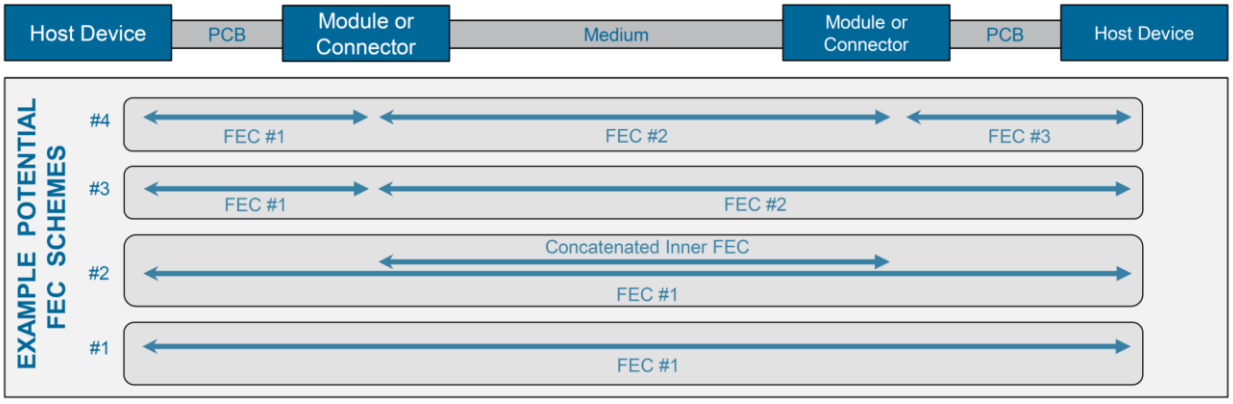


Figure 4-11. *FEC Schemes for 800Gb/s and 1.6Tb/s Ethernet rates.*

The key performance factors to evaluate a FEC code are: (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

- The coding gain, where a new tradeoff is needed on the additional overhead of the transceivers and the links, with the new error correct performance applied.
- The latency, where the longer FEC codewords and the higher error correction performance leads to the higher latency. The latency may be lower in some degree, using higher parallelism in circuit implementations.
- The area and power, which is mostly represented by the hardware implementation complexity, and they are generally increased by using higher parallelism in circuit implementations. A new tradeoff is needed between the area/power and the latency, which is mostly an implementation challenge and does not affect the interoperability.

IEEE 802.3 B400G TF considers the following FEC schemes for 800Gb/s and 1.6Tb/s transmission rates: (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

- the End-to-End,
- the Encapsulated or Concatenated, and
- the Segmented.

The FEC schemes for 800Gb/s and 1.6Tb/s are summarized in **Table 4-2**.

FEC Scheme	AUI Interface	Optical Lanes	Reach
End-to-End (E2E)	800Gb/s Ethernet: 8x100Gb/s & 4x200Gb/s 1.6Tb/s Ethernet: 16x100Gb/s & 8x200Gb/s	800Gb/s Ethernet: 8x100Gb/s 1.6Tb/s Ethernet: 16x100Gb/s	50m-500m
Encapsulated or Concatenated		800Gb/s Ethernet: 4x200Gb/s 1.6Tb/s Ethernet: 8x200Gb/s	50m-2km
Segmented		800Gb/s Ethernet: 1x800Gb/s 1.6Tb/s Ethernet: 2x800Gb/s & 1x1600Gb/s	10km+

Table 4-2. *Assumptions of Electrical/Optical Lanes for FEC Analysis.*

4.3.1 End-to-End FEC Scheme

The End-to-End FEC scheme uses the current specifications of IEEE 802.3bs/cu/ck for 100Gb/s and 200Gb/s per lane transmission, providing 400Gb/s by using multiple lanes (4 and 2 lanes, respectively). End-to-End FEC scheme uses one only FEC code between the transmitter and the receiver side, where at the latter one FEC decoder corrects all errors caused by the whole link, such as the two AUI instances and the optical cable link, as shown in **Figure 4-12**. (Wang & He, 2022)

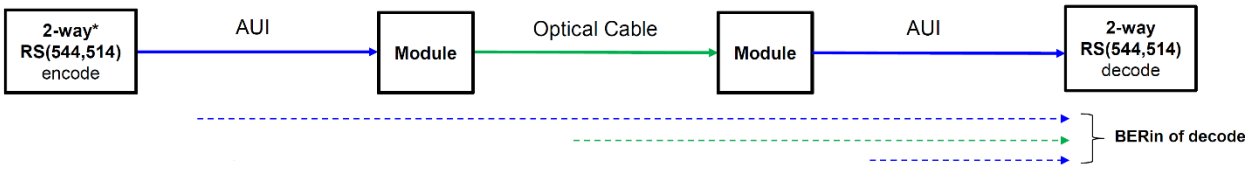


Figure 4-12. *End-to-End FEC Scheme.*

In End-to-End FEC scheme can be fitted any FEC code in principle, such as the RS-FEC, the BCH, the Hamming, the Concatenated FEC (CFEC), the Open FEC (OFEC), etc. However, the RS(544,514) can be reused from IEEE 802.3bs/cu/ck to enable backward compatibility to 10Gb/s per lane AUI and PHY. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

The FEC sublayer allocation and the End-to-End FEC scheme for 800Gb/s and 1.6Tb/s in ISO OSI reference model are shown in **Figure 4-13**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

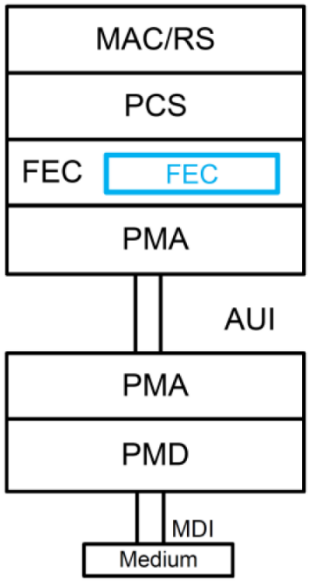


Figure 4-13. *FEC Sublayer Allocation for End-to-End FEC Scheme.*

4.3.2 Encapsulated or Concatenated FEC Scheme

The Encapsulated or Concatenated FEC scheme introduced for first time by G. David Forney's paper "Concatenated Codes" in 1965. The Encapsulated or Concatenated FEC scheme comprises both an outer and an inner code, which are operating together to correct the errors in a message as shown in **Figure 4-14**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

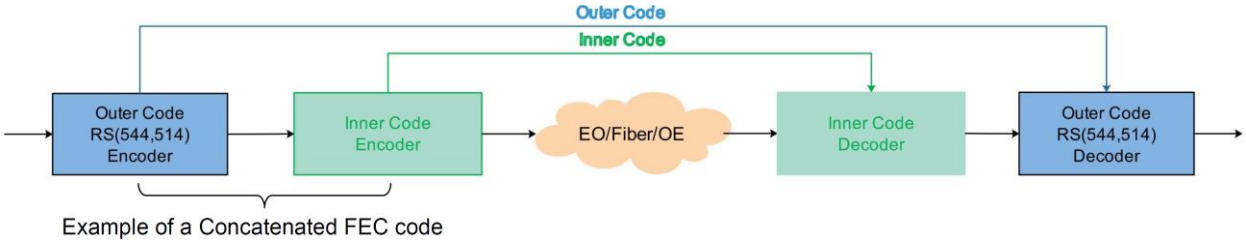


Figure 4-14. *Concatenated FEC Code.*

In general, the outer code has a higher coding gain than the inner code as specified in IEEE 802.3cw (400GBASE-ZR), where a Staircase FEC (SFEC) code is used as the outer code and a Hamming code is used as the inner code. An example of the Encapsulated or Concatenated FEC scheme is the ITU.T G.975 - Forward Error Correction for High Bit-rate DWDM Submarine Systems, which uses the RS(544,514) as

the outer code and the BCH/Hamming as the inner code. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

The Encapsulated or Concatenated FEC scheme uses one only Concatenated FEC code, which at the receive side of the inner FEC decoder, it corrects the errors caused only by the inner link, such as the optical cable link and at the receive side of the outer FEC decoder, it corrects the errors caused by the whole link, including both the two AUI instances and the left-over errors from the optical cable link, as shown in **Figure 4-15**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

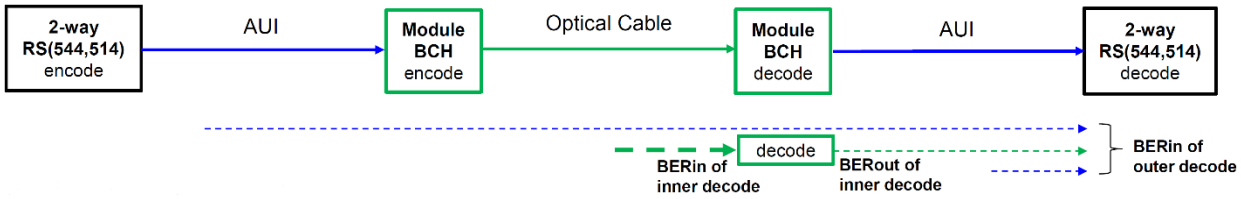


Figure 4-15. *Encapsulated or Concatenated FEC Scheme.*

In Encapsulated or Concatenated FEC scheme, the RS(544,514) FEC code can be used as specified in IEEE 802.3bs/cu/ck as the outer code for both the random and the burst error tolerance and the Soft Decision/Hard Decision (SD/HD) BCH/Hamming code as the inner code. This scheme can provide a lower raw Bit Error Rate (BER) to enable 100Gb/s and 200Gb/s per lane optical link. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

The FEC sublayer allocation and the Encapsulated or Concatenated FEC scheme for 800Gb/s and 1.6Tb/s in ISO OSI reference model are shown in **Figure 4-16**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

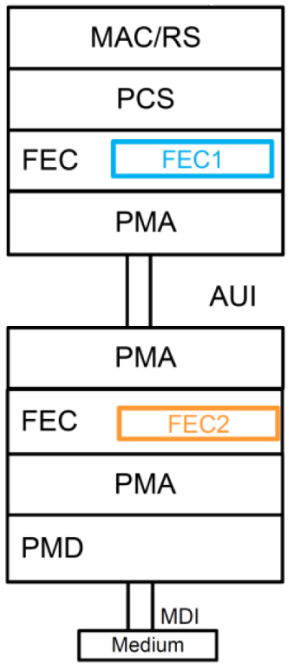


Figure 4-16. *FEC Sublayer Allocation for Encapsulated or Concatenated FEC Scheme.*

where:

FEC1 is a 2-way interleave RS(544,514) code, and

FEC2 can be a new inner BCH/Hamming code or an OFEC/CFEC code.

Further consideration is needed when selecting the FEC codes for FEC1 as the outer code and FEC2 as the inner code, as follows: (Wang & He, Analysis on FEC1 in Archtecture of 800G/1.6TbE, 2022)

- A Hard-Decision RS(544,514) code can be used for the FEC1 as specified in IEEE 802.3bs.
- FEC2 can manage BER requirements of the optical cable link, e.g. the 200Gb/s per lane Intensity Modulation - Direct Detection (IM-DD) link has a pre-FEC BER of approximately 2x10-3 by using a Soft-Decision RS(544,514) as the outer code and a pre-FEC BER of approximately 1x10-3 by using a Hard-Decision BCH (t=1 or 2) as the inner code.
- A pre-FEC BER of 1x10-5 for 200Gb/s per lane AUI can allocate a 0.1dBo FEC coding gain as specified in IEEE 802.3bs.
- Interleaving FEC1 need to be considered for non-Poisson distributed errors from FEC2 output with an as low as possible latency.

- A higher overhead for the BCH/Hamming code as the inner code is less challenging than a further lower raw BER at the optical cable links.

A proposal for the FEC1 as specified in IEEE 802.3bs uses two RS(544,514) FEC codewords interleaving for both 800Gb/s and 1.6Tb/s transmission rates, with eight and sixteen FEC lanes respectively, corresponding to 100Gb/s per lane AUI objective, as shown in **Figure 4-17**. (Wang & He, Analysis on FEC1 in Archtecture of 800G/1.6TbE, 2022)

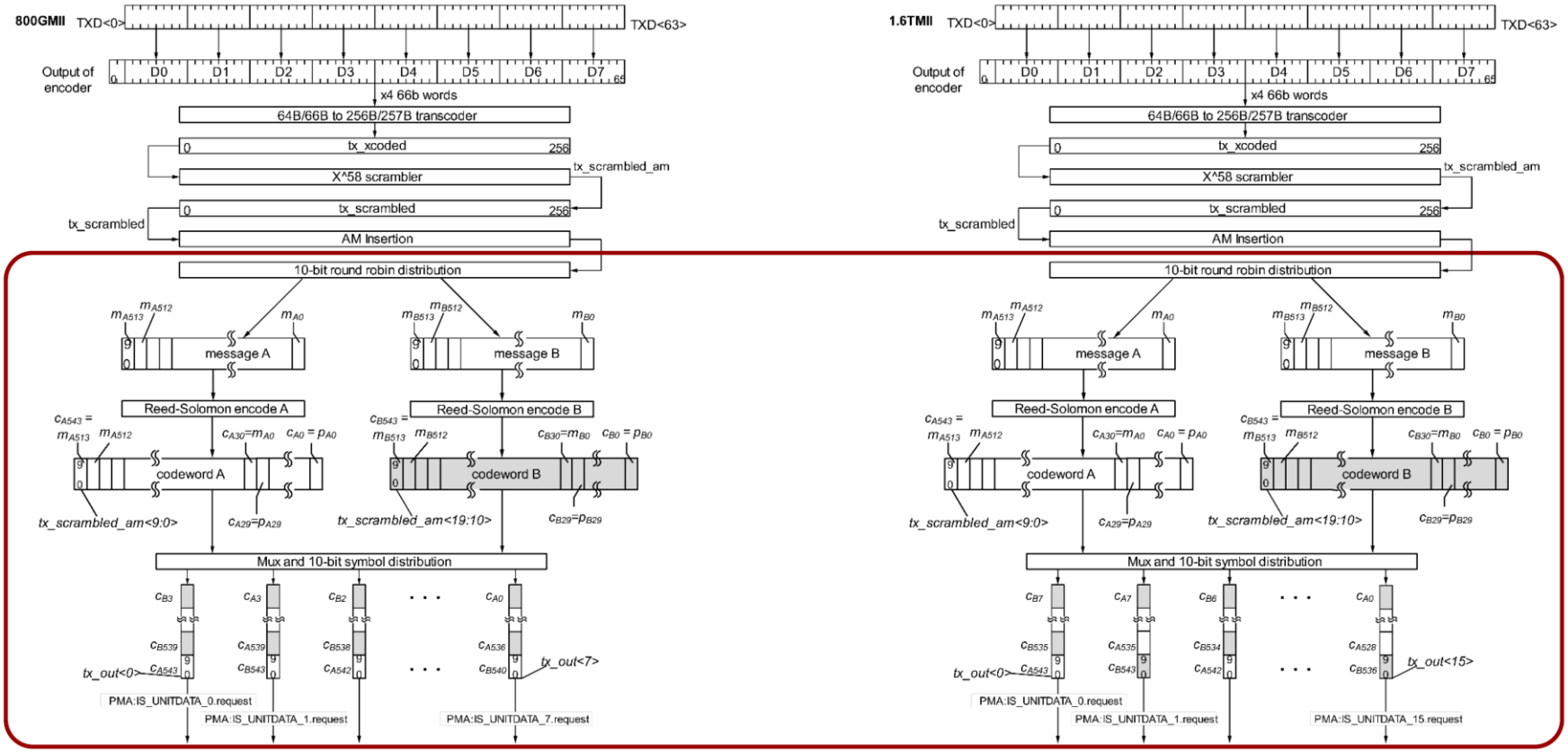


Figure 4-17. *Proposed FEC1 Code.*

The proposed FEC1 code characteristics are summarized in **Table 4-3**. (Wang & He, Analysis on FEC1 in Archtecture of 800G/1.6TbE, 2022)

Characteristics	800Gb/s Ethernet	1.6Tb/s Ethernet
Interleave	2 x FEC codeword	2 x FEC codeword
FEC lanes	8 FEC lanes @ 100Gb/s	16 FEC lanes @ 100Gb/s
Alignment Marker	8	16
PMA	1:1 for 100Gb/s per lane 2:1 for 200Gb/s per lane 4:1 for 400Gb/s per lane	1:1 for 100Gb/s per lane 2:1 for 200Gb/s per lane 4:1 for 400Gb/s per lane

Table 4-3. *Proposed FEC1 Characteristics.*

4.3.3 Segmented FEC Scheme

The Segmented FEC scheme uses three independent FEC codes for each one of the three segments (parts of a link). At the receive side of each segment - the two AUI segments and the optical cable link – the FEC decoder corrects only the errors caused by its corresponding segment as shown in **Figure 4-18**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

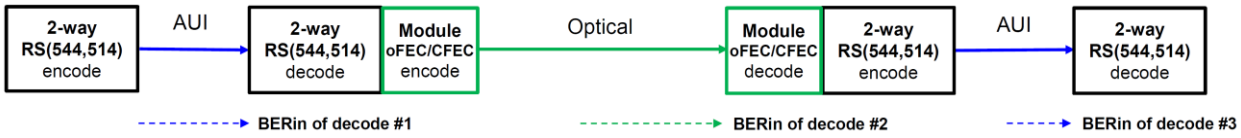


Figure 4-18. *Segmented FEC Scheme.*

The FEC decoder can reuse the RS(544,514) code specified in IEEE 802.3bs/cu/ck for 100Gb/s per lane AUIs, while for 200Gb/s per lane optical link the FEC code can be defined based on the target BER. In case the optical cable link operates at a BER of approximately 2x10-3, a higher performance FEC code with a coding gain greater than 8dB should be used, comparing with the approximately 6.4dB coding gain of the RS(544,514) code. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

The FEC sublayer allocation and the Segmented FEC scheme for 800Gb/s and 1.6Tb/s in ISO OSI reference model are shown in **Figure 4-19**. (Wang & He, FEC Code and Scheme Observation in 800G/1.6TbE, 2022)

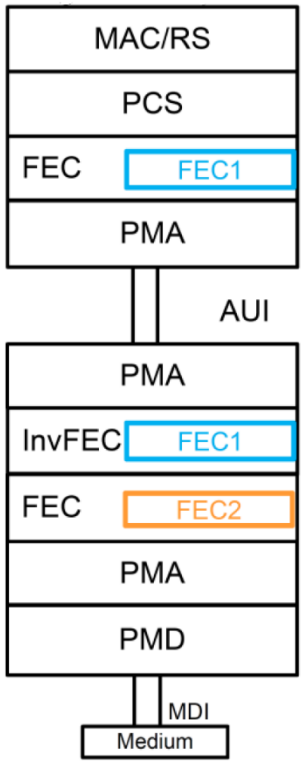


Figure 4-19. *FEC Sublayer Allocation for Segmented FEC Scheme.*

where:

FEC1 is a 2-way interleave RS(544,514) code, and

FEC2 can be a new inner BCH/Hamming code or a OFEC/CFEC code.

The Segmented FEC scheme allows the decoupling of FEC1 and FEC2, where there are no dependencies between them on each particular FEC code selection. The FEC code of each segment has a deterministic pre-FEC BER limitation for a target post-FEC BER. (Wang & He, Analysis on FEC1 in Archtecture of 800G/1.6TbE, 2022)

The FEC schemes which support the proposed BER objectives are summarized in ***Table 4-4***: (Wang, He, & Ren, BER objective for Beyond 400GbE, 2021)

FEC Approach	FEC Algorithm Example *	Pre-FEC BER (Rough estimated)	Net Coding Gain (Assume 1×10^{-14} BER Objective)	Optical PMD Reach Example*
End-to-End (E2E)	RS (544,514) (Hard Decision)	≈$2.2\text{x}10^{-4}$	≈6.6dB	50m - 500m
Encapsulated or Concatenated	Outer code: RS (544,514) (Hard Decision) Inner code: BCH, t=1 or 2 (Soft/Hard Decision)	≈$2\text{x}10^{-3}$ ≈10^{-3}	≈8.4dB ≈7.7dB	50m - 2km+
Segmented	CFEC and/or Product Code	≈10^{-2}	≈10.4dB	10km+
*FEC approaches are independent to PMD solution, IM-DD/Coherent, and different reach.				

Table 4-4. *FEC approaches to support proposed BER objectives.*

4.3.4 Considerations between FEC and DSP

It is important and efficient to improve the channel and use an advance Digital Signal Processing (DSP) rather than proceeding to a stronger FEC coding. The FEC coding technologies are much less efficient using encoding schemes beyond RS(544,514) FEC code, especially in the End-to-End FEC scheme. The RS(576,514) FEC code with Maximum Likelihood Sequence Equalizer (MLSE) may give the best performance. (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

In case a more powerful or stronger FEC coding must be used to address the performance concern, it is necessary to consider a mutual exclusive relationship of FEC coding with the advance DSP technologies and an architecture to support the End-to-End, the Encapsulated or Concatenated and the Segmented FEC schemes. (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

The relationship of FEC coding and DSP is shown in **Figure 4-20**, that is the relation between the equalizers (Feed-Forward Equalizer (FFE), Decision Feedback Equalizer (DFE) and MLSE) and the FEC codes (Hard-Decision FEC (HD-FEC) and Soft-Decision

FEC (SD-FEC)). (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

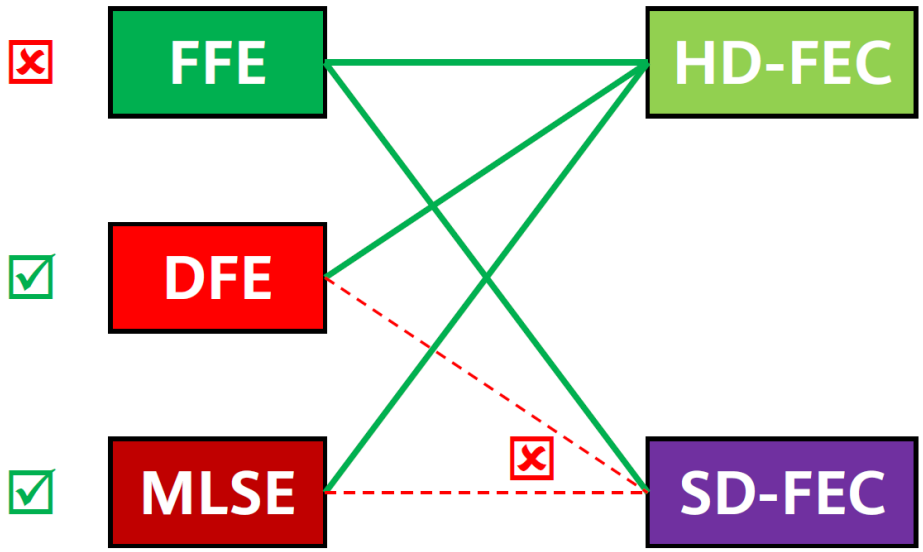


Figure 4-20. *Relationship of FEC and DSP.*

As illustrated in **Figure 4-20**, the HD-FEC has the best compatibility with all the DSP algorithms.

However, the SD-FEC code is difficult to be compatible with the DFE and the MLSE non-linear equalizers. The DFE has a much better performance than FFE, since the propagated error leads to wrong information to the SD-FEC decoder and is harmful to the decoding process. The MSLE requires soft-output algorithms, such as the SOft-decision Viterbi Algorithm (SOVA), and the Bahl, Cocke, Jelinek, Raviv (BCJR) with an unknown performance, which are not cost-effective and may be un-affordable for the real applications. (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

The DFE is essential for 100Gb/s as it has already used as a reference receiver in 100Gb/s electrical links with a DFE coefficient of 0.85. There is no evidence showing that DFE is sufficient for 200Gb/s optics. (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

MLSE is more efficient than SD-FEC, so as it has no overclocking and a maximum of 2dB net-coding gain. (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

The new approach for the DSP architecture is shown in **Figure 4-21**. (LU & Zhuang, DSP and FEC Considerations for 800GbE and 1.6TbE, 2022)

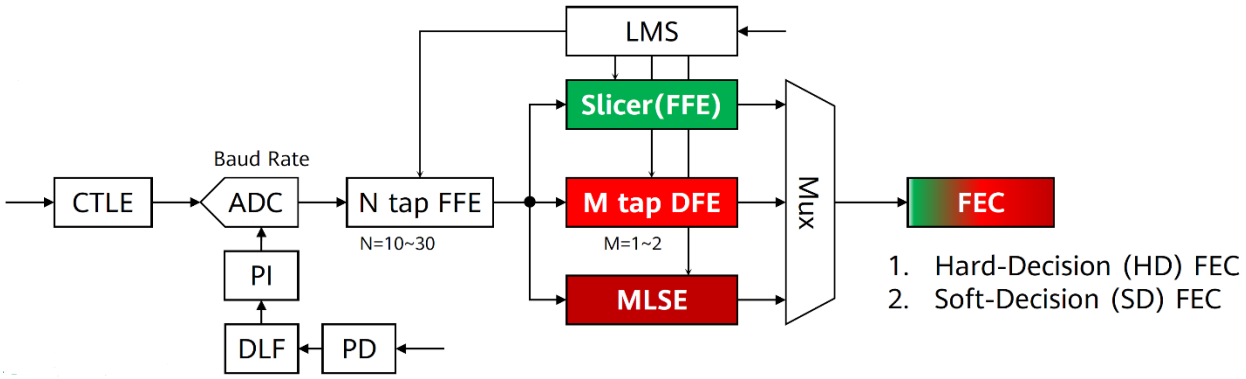


Figure 4-21. *New Approach of DSP Architecture.*

where the peripheral functional blocks are the following:

- Continuous Time Linear Equalizer (CTLE).
- Analogue to Digital Converter (ADC).
- Phase Interpolator (PI).
- Delay Line Filter (DLF).
- PhotoDetector (PD).
- Least-Mean Square (LMS).
- Multiplexer (MUX).

4.4 Physical Medium Attachment (PMA) Sublayer

The PMA operates as defined in Section **3.5**, with the exception that there are 32 PCS lanes and only 4:1-bit multiplexing (muxing) is performed. The PMA has complete freedom on multiplexing any PCS lanes together, PCS lanes are in no way restricted on their location on any PMD or AUI lane. (G, G, & O, 2021)

So, when is needed to multiplex logical lanes together for 800Gb/s and 1.6Tb/s, several options have been reused or re-identified, such as: (Gustlin, Pepper, Szczepanek, Trowbridge, & Wang, 2014)

- Bit muxing, used in IEEE 802.3ba,
- FEC Orthogonal Multiplexing,
- Block muxing.

Bit muxing is preferred because it enables protocol agnostic optical module and friendly reuse in non-Ethernet interconnect area. However, block muxing has a better performance in FEC for burst error than bit muxing and it is a protocol aware optical module as delimiter block boundary. (Wang, He, & Ren, Technical Feasibility of Logic Layer to Support Rate Objective, 2021)

Reusing 400Gb/s PMAs for 800Gb/s and 1.6Tb/s, the PMA in PCS is 16:16 or 16:8 with symbol-mux/demux and PMA[16:4], PMA[16:8] or PMA[8:4] in module is bit-mux/demux. Then group the FEC lanes from 4 different RS-FECs and do bit-mux/demux in optical module, as illustrated in **Figure 4-22**. In this case, there is no specific lane sequence and skew constraint in each bit-mux/demux group. If burst error happens in each group, it breaks to shorter errors, then it is distributed to different sub RS-FEC codewords. (Wang, Wang, & Yang, Technical Feasibility of Bit multiplexing in 400GbE PMA, 2014)

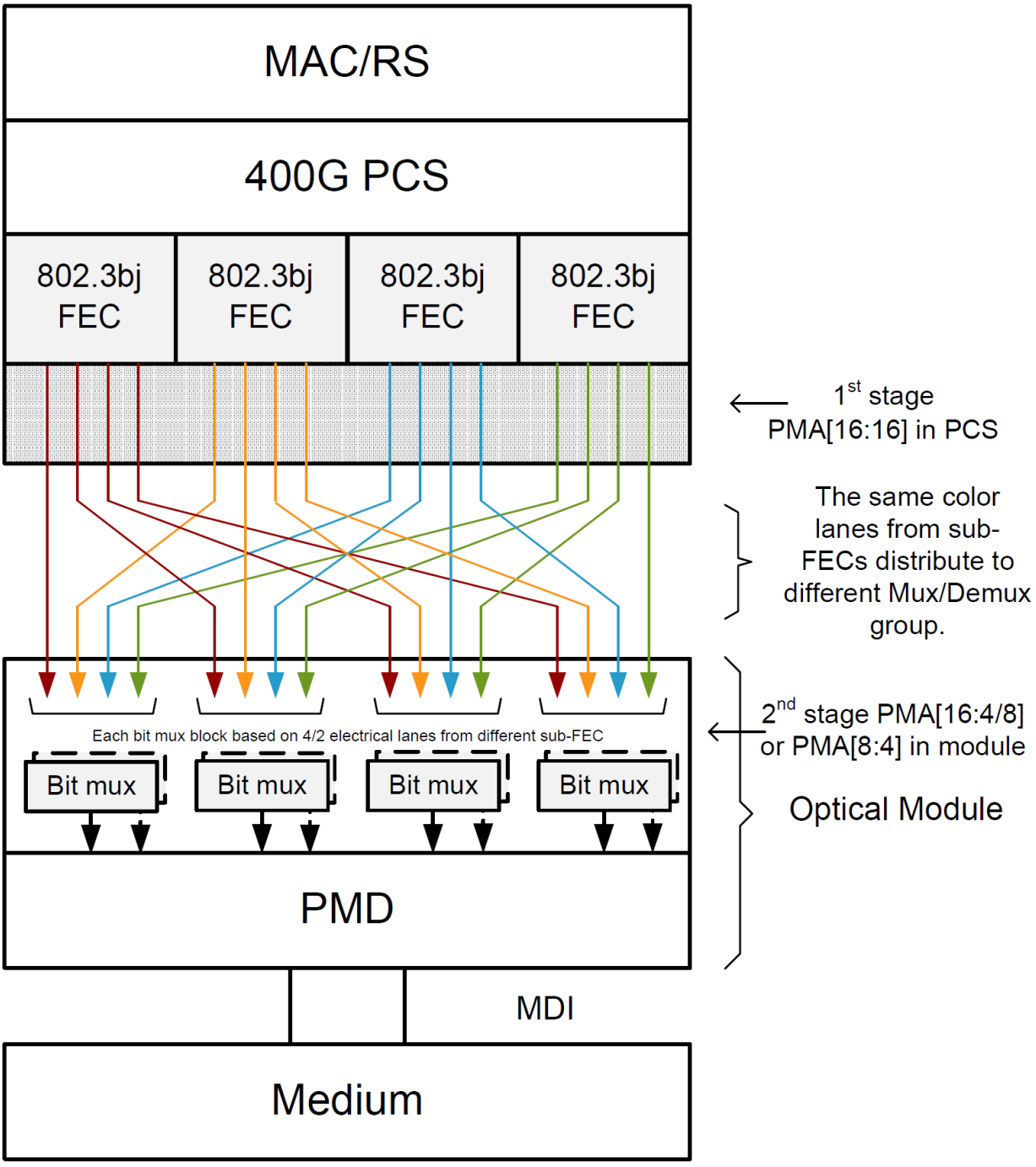


Figure 4-22. *Improved PCS/PMA architecture with bit-mux.*

The lanes alignment and reorder mechanism in transmit and receive side across all lanes is illustrated in **Figure 4-23** and is according the 100Gb/s Ethernet. (Wang, Wang, & Yang, Technical Feasibility of Bit multiplexing in 400GbE PMA, 2014)

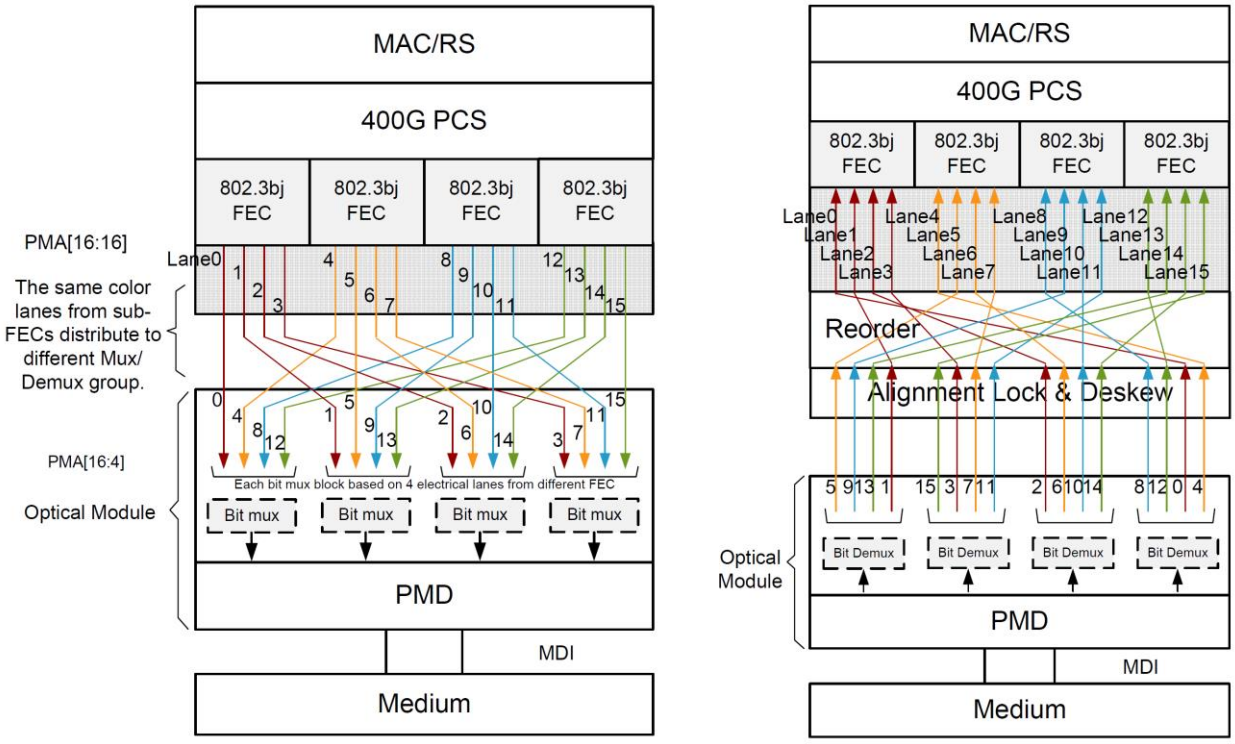


Figure 4-23. *FEC lanes alignment and reorder mechanism.*

100GAUI-n, 200GAUI-n and 400GAUI-n use either NZR or PAM4 physical instantiations. 800Gb/s and 1.6Tb/s is proposed to use either PAM4, PAM6 or PAM16 physical instantiations.

From NRZ to PAM4, the data rate is doubled with baud rate unchanged, as shown in **Figure 4-24**. The SNR loss is compensated by RS(544, 514) FEC. The upgrade from PAM4 to PAM16 does not work, because the SNR loss is too high and the FEC is less cost-effective. In this case, one possible way to double the data rate of differential pairs and keep the baud rate unchanged is to use Single-Ended (SE) signaling. The intra-pair crosstalk can be canceled by MIMO algorithm which is mature. 200Gb/s optical links need 100GBd PAM4, because the optical links are native "Single-Ended" system. (LU & Zhuang, 2021)

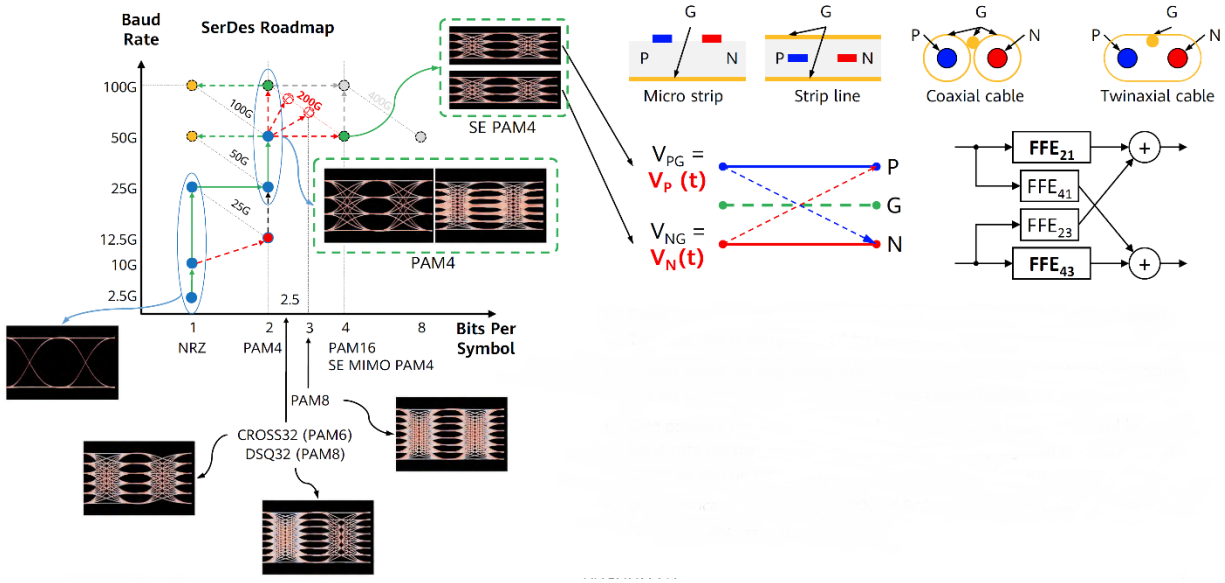


Figure 4-24. *Beyond 100Gb/s SerDes (PMA).*

200Gb/s PAM signaling constellations are shown in **Figure 4-25**:

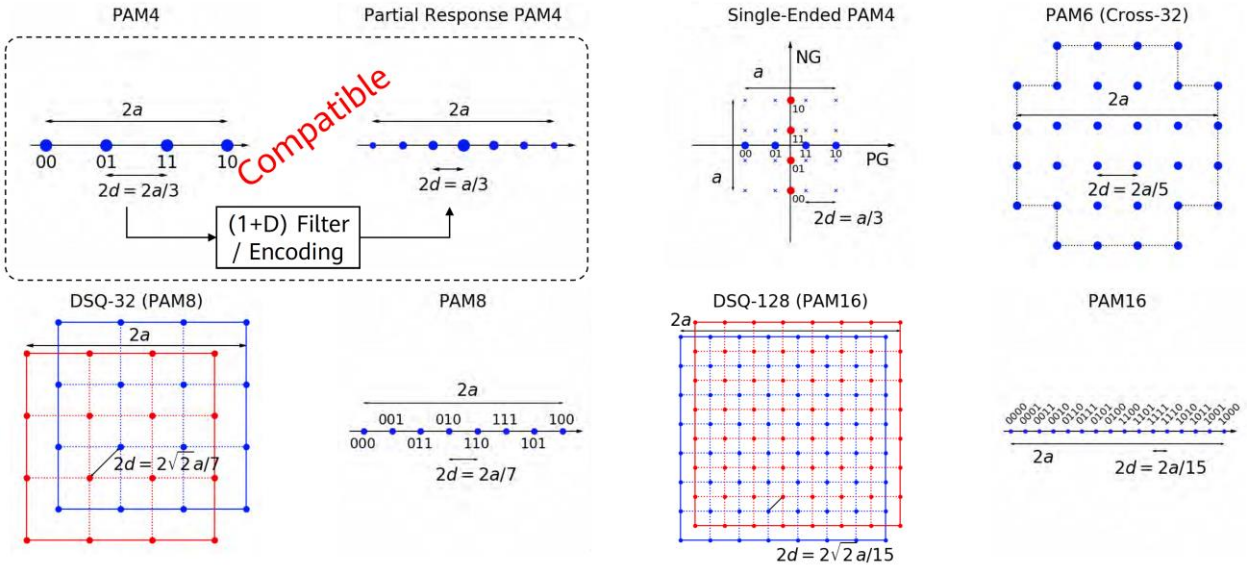


Figure 4-25. *200G PAM signaling constellations.*

200Gb/s PAM signaling Eye diagrams are shown in **Figure 4-26**:

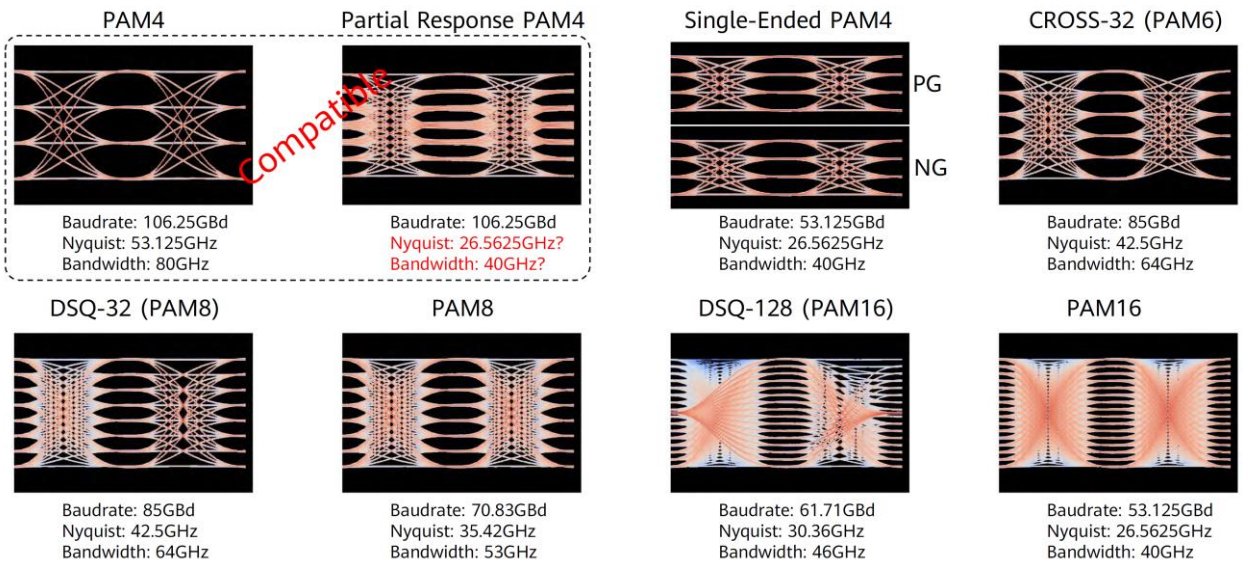
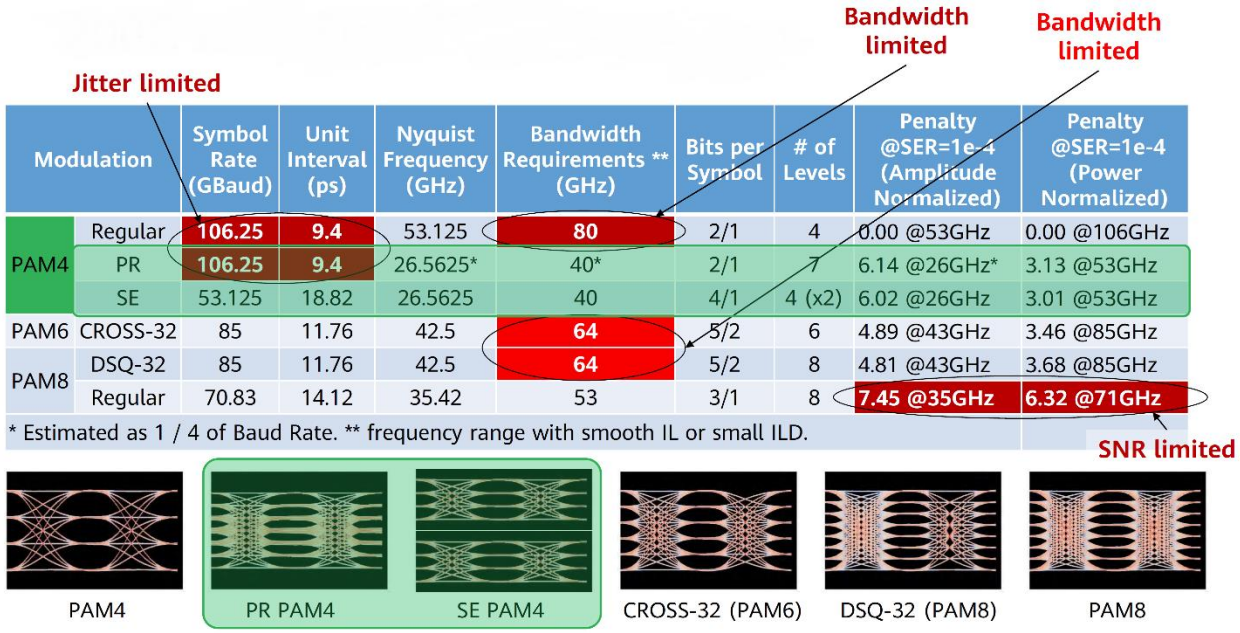


Figure 4-26. *200G PAM signaling eye diagrams.*

In **Figure 4-27** is shown a comparison of 200Gb/s PAM signaling schemes:



Modulation		Symbol Rate (GBaud)	Unit Interval (ps)	Nyquist Frequency (GHz)	Bandwidth Requirements ** (GHz)	Bits per Symbol	# of Levels	Penalty @SER=1e-4 (Amplitude Normalized)	Penalty @SER=1e-4 (Power Normalized)
PAM4	Regular	106.25	9.4	53.125	80	2/1	4	0.00 @53GHz	0.00 @106GHz
	PR	106.25	9.4	26.5625*	40*	2/1	4	6.14 @26GHz*	3.13 @53GHz
	SE	53.125	18.82	26.5625	40	4/1	4 (x2)	6.02 @26GHz	3.01 @53GHz
PAM6	CROSS-32	85	11.76	42.5	64	5/2	6	4.89 @43GHz	3.46 @85GHz
PAM8	DSQ-32	85	11.76	42.5	64	5/2	8	4.81 @43GHz	3.68 @85GHz
	Regular	70.83	14.12	35.42	53	3/1	8	7.45 @35GHz	6.32 @71GHz

* Estimated as 1 / 4 of Baud Rate. ** frequency range with smooth IL or small ILD.

Figure 4-27. *Comparison of 200Gb/s PAM signaling schemes.*

The use of a PAM4 Digital Signal Processor (DSP) for Direct Detect (DD), 100Gbps, has already been used in a 28nm CMOS. In the latest generation, a 400Gbps module design uses a QSFP-DD (Quad Small Form-Factor Pluggable - Double Density) form factor with a 7nm DSP, dissipating 4W/100Gbps typically, at 16W for the 400Base-ZR module. (Nagarajan, Lyubomirsky, & Agazzi, 2021)

The DSP Application-Specific Integrated Circuit (ASIC) power is nominally about 50% of the module. Since the optical modules typically operate from a 3.3 V supply, and most high-end ASICs operate at supply voltages less than 1V, there is a power conversion loss (segment labeled “power overhead”) that is between 10% and 15%. For the DSP power to be nominally 50% of the total module power, the complexity of the DSP and optics need to scale with applications. For coherent applications, the laser is full C-band tunable, and its power consumption is in the slice labeled “tunable laser”. In the 100Gbps PAM4 Dense Wavelength Division Multiplexing (DWDM) applications, fixed wavelength lasers in the C-band are used for lower power consumption. (Nagarajan, Lyubomirsky, & Agazzi, 2021)

The slice labeled “optics + electronics” has the modulator driver, TransImpedance Amplifier (TIA), and the integrated silicon photonics chip which has the high-speed Mach Zehnder modulators and photodetectors. (Nagarajan, Lyubomirsky, & Agazzi, 2021)

4.5 Physical Medium Dependent (PMD) Sublayer

The IEEE B400G TF starts considering the PMDs for 800Gb/s and 1.6Tb/s Ethernet to address copper, MMF, and SMF specifications in both IEEE P802.3df and IEEE 802.3dj standards. For 800Gb/s and 1.6Tb/s Ethernet the proposed PMD types are:

- MMF (current trend up to 100m),
- SMF - parallel fiber approach (current trend up to 500m),
- SMF - duplex fiber approach (2km to 40km),
- DWDM Systems.

For 800Gb/s with 8 lanes of 100Gb/s (IEEE P802.3df), based on IEEE 802.3ck, the proposed PMD objectives are: (LU & ZHUANG, Considerations on beyond 100G per lane electrical objectives, 2021)

- An eight-lane 800Gb/s AUI for C2M applications, compatible with PMDs based on 100Gb/s per lane optical signaling.
- An eight-lane 800Gb/s AUI for chip-to-chip (C2C) applications.

For 800Gb/s with 4 lanes of 200Gb/s (IEEE P802.3dj), the proposed PMD objectives are: (LU & ZHUANG, Considerations on beyond 100G per lane electrical objectives, 2021)

- A four-lane 800Gb/s AUI for C2M applications, compatible with PMDs based on 200Gb/s per lane optical signaling.
- A four-lane 800Gb/s AUI for C2C applications.
- A four-lane 800Gb/s AUI for chip-to-near-package-optics applications, compatible with PMDs based on 200Gb/s per lane optical signaling or merged with C2C and C2M objectives.

For 1.6Tb/s with 8 lanes of 200Gb/s (IEEE P802.3dj), the proposed PMD objectives are: (LU & ZHUANG, Considerations on beyond 100G per lane electrical objectives, 2021)

- An eight-lane 1.6Tb/s AUI for C2M applications, compatible with PMDs based on 200Gb/s per lane optical signaling.
- An eight-lane 1.6Tb/s AUI for C2C applications.
- An eight-lane 1.6Tb/s AUI for chip-to-near-package-optics applications, compatible with PMDs based on 200Gb/s per lane optical signaling or merged with Chip-to-Chip (C2C) and Chip-to-Module (C2M) objectives.

4.6 Fiber Optic Cabling Model

The fiber optic cabling model for 800Gb/s and 1.6Tb/s will remain the same as in **Figure 2-23**, **Figure 2-29**, **Figure 2-49**, **Figure 3-22**, **Figure 3-29**, **Figure 3-32** and **Figure 3-38** for DR4, FR4, LR4, FR8, LR8, SR4, SR10 and SR16, respectively.

The channel IL will be defined in the next step of each standard (IEEE P802.3df and IEEEP802.3dj) for 200 Gb/s, 400 Gb/s, 800 Gb/s, and 1.6 Tb/s Ethernet Task Force. In any case the channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion meet the specifications.

The specifications of fiber optic cabling for 800Gb/s and 1.6Tb/s will be defined in the next steps of the IEEE B400G TF.

Annex A. 200Gb/s and 400Gb/s PMA sublayer partitioning examples

This Annex provides various partitioning examples. Partitioning guidelines and MMD numbering conventions are described in Section **3.5**.

A.1 Partitioning example supporting 400GBASE-SR16

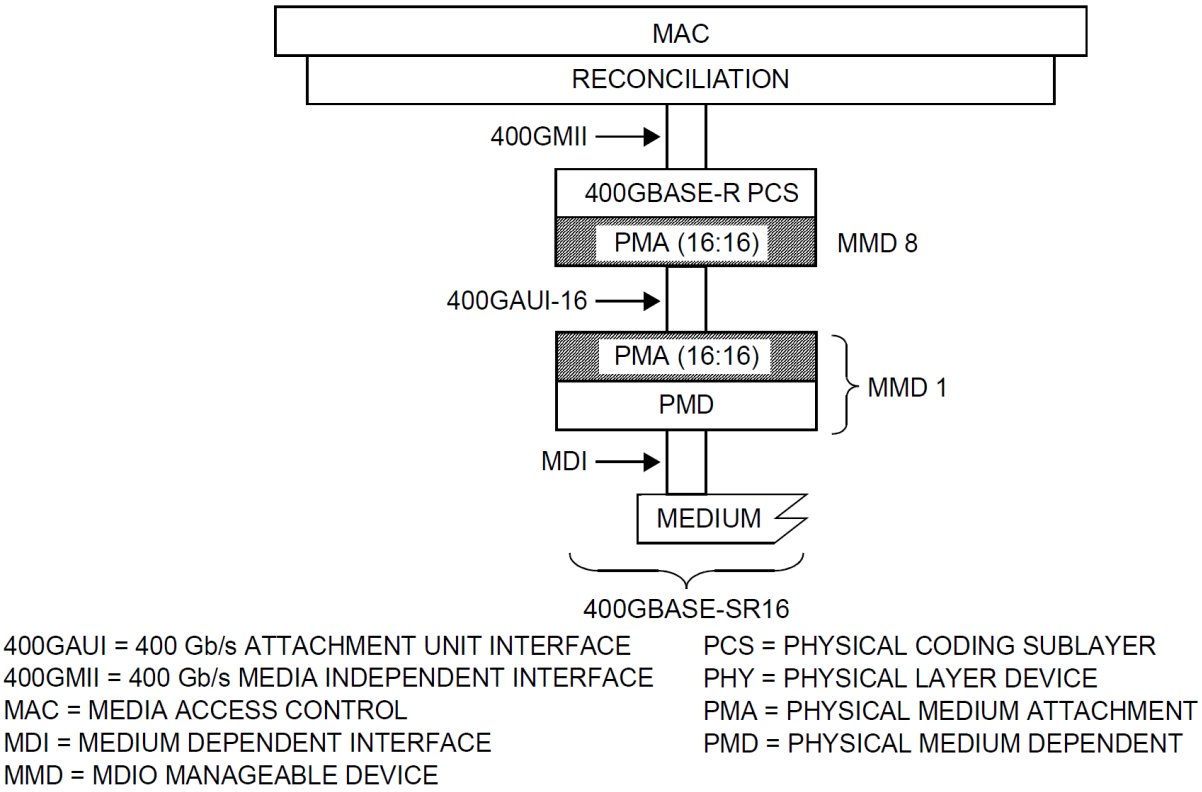


Figure A. 1. *Example 400GBASE-SR16 PMA layering with single 400GAUI-16 chip-to-module interface.*

A.2 Partitioning example supporting 200GBASE-DR4/FR4/LR4 and 400GBASE-FR8/LR8

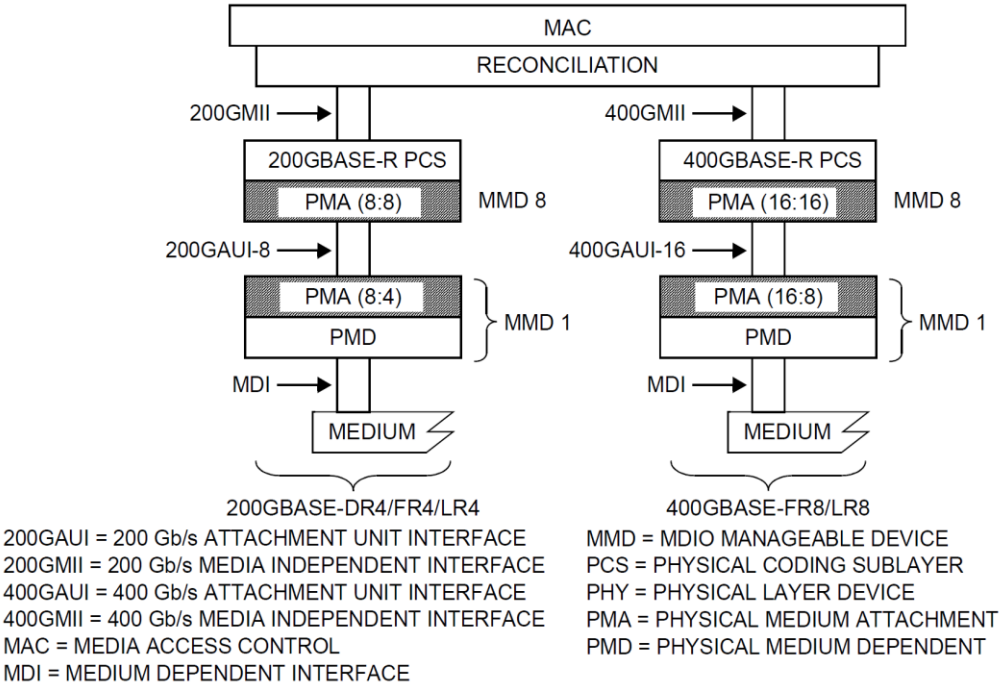


Figure A. 2. *Example 200GBASE-DR4/FR4/LR4 or 400GBASE-FR8/LR8 PMA layering with single 200GAUI-8 or 400GAUI-16 chip-to-module interface.*

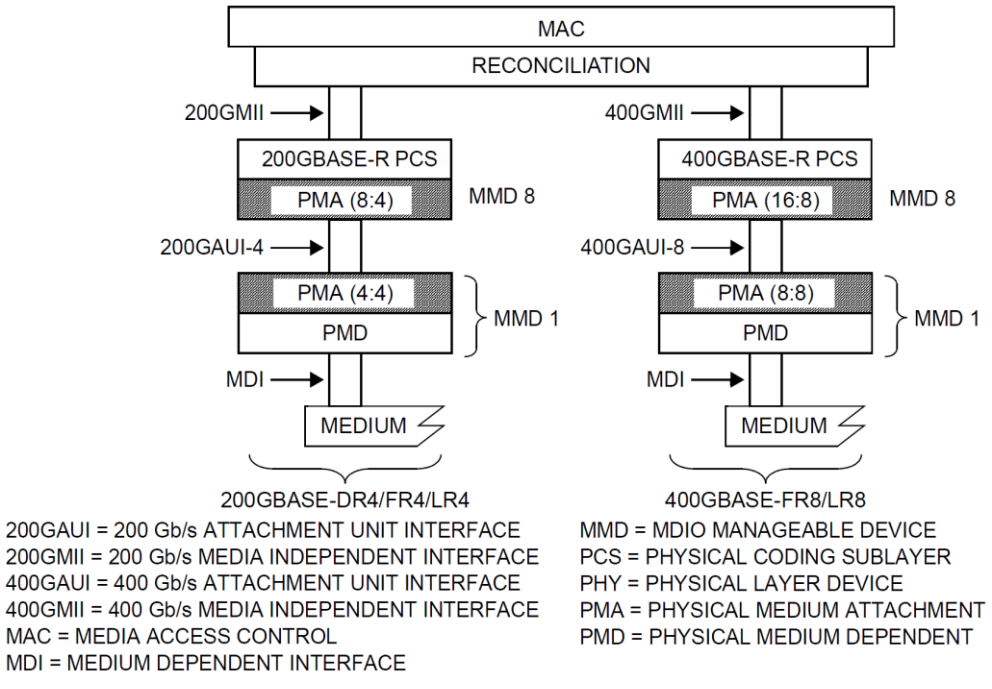


Figure A. 3. *Example 200GBASE-DR4/FR4/LR4 or 400GBASE-FR8/LR8 PMA layering with single 200GAUI-4 or 400GAUI-8 chip-to-module interface.*

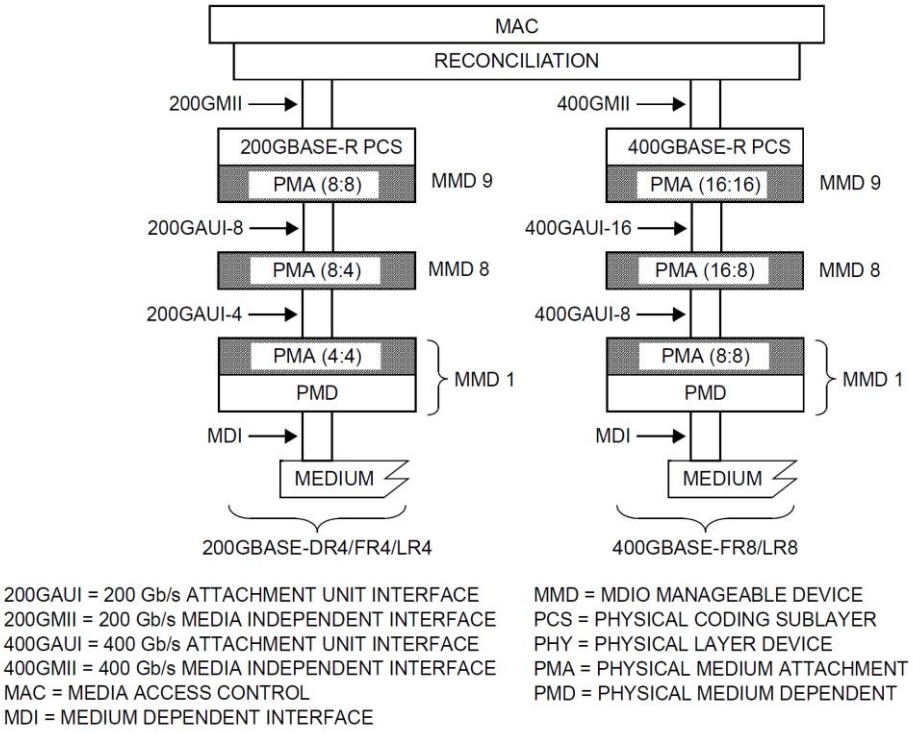


Figure A. 4. *Example 200GBASE-DR4/FR4/LR4 or 400GBASE-FR8/LR8 PMA layering with 200GAUI-8 or 400GAUI-16 chip-to-chip and 200GAUI-4 or 400GAUI-8 chip-to-module interfaces.*

A.3 Partitioning example supporting 400GBASE-DR4

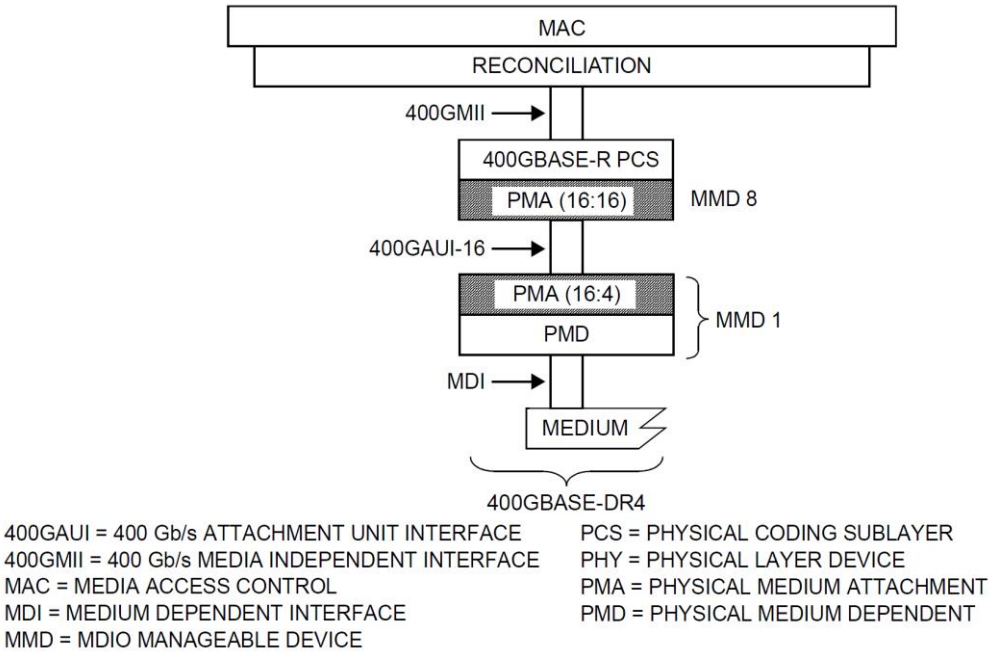
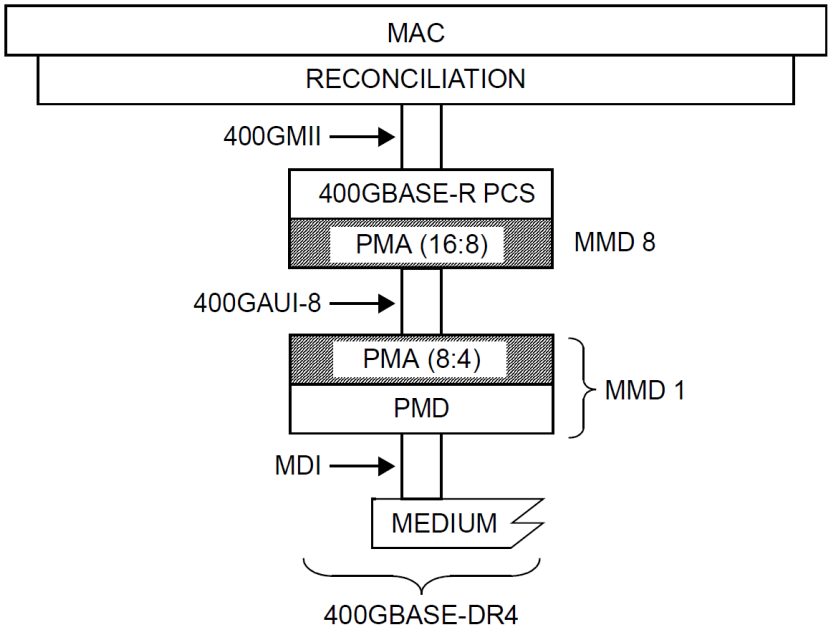


Figure A. 5. *Example 400GBASE-DR4 PMA layering with single 400GAUI-16 chip-to-module interface.*



400GAUI = 400 Gb/s ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure A. 6. *Example 400GBASE-DR4 PMA layering with single 400GAUI-8 chip-to-module interface.*

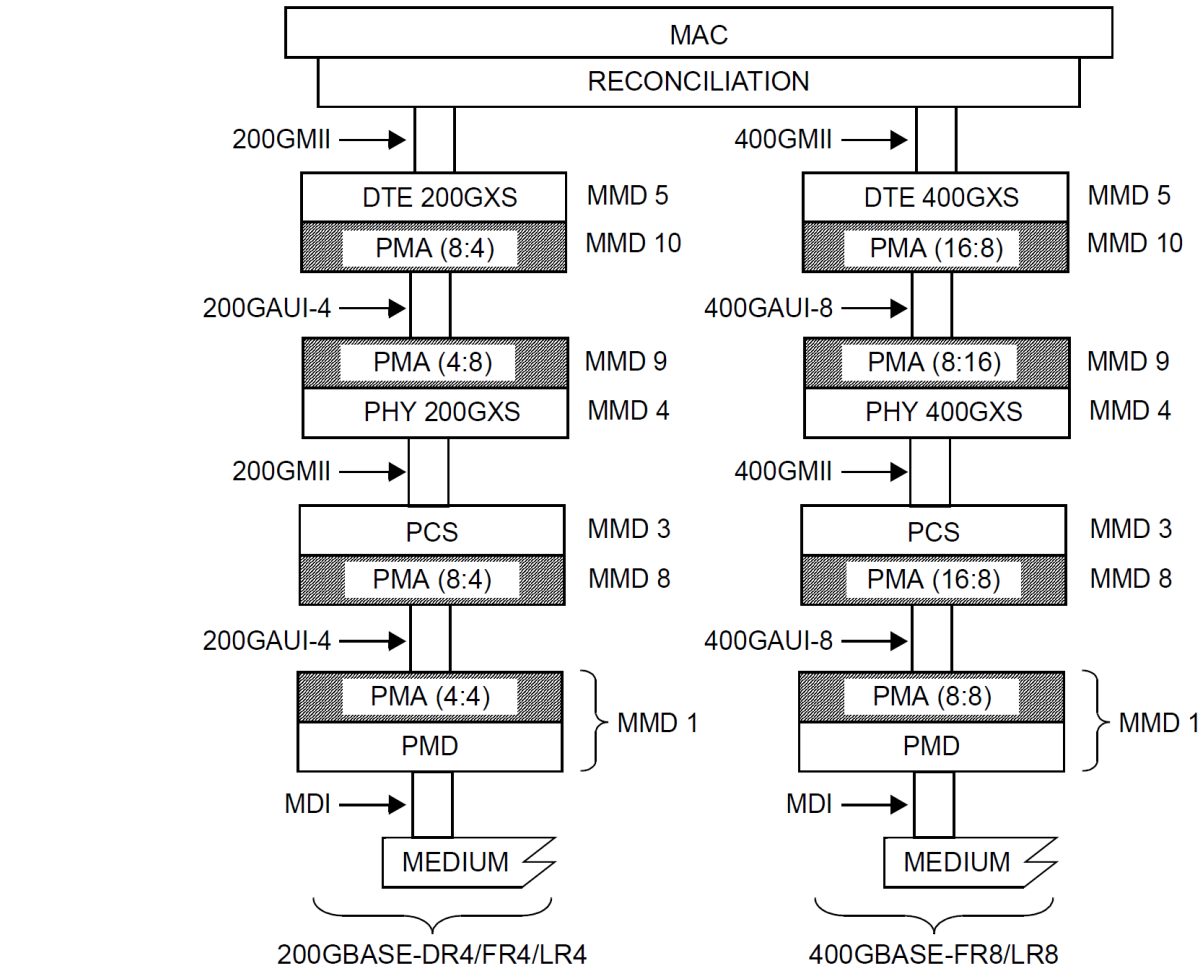
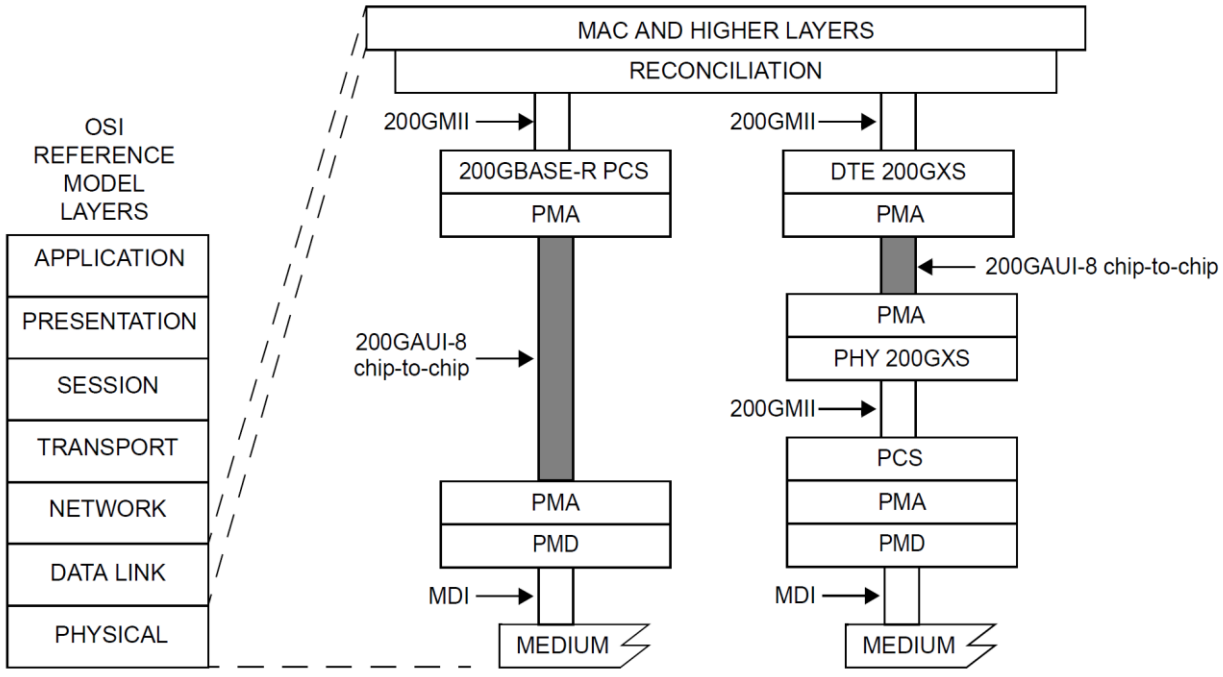


Figure A. 7. *Example 200GBASE-DR4/FR4/LR4 and 400GBASE-FR8/LR8 PMA layering with 200GXS, 400GXS, and two 200GAUI-4, 400GAUI-8 interfaces.*

Annex B. Chip-to-chip 200 Gb/s eight-lane Attachment Unit Interface (200GAUI-8 C2C) and 400 Gb/s sixteen-lane Attachment Unit Interface (400GAUI-16 C2C)

This annex defines the functional for the optional chip-to-chip 200Gb/s eight-lane AUI (200GAUI-8 C2C) and 400 Gb/s sixteen-lane AUI (400GAUI-16 C2C). **Figure B. 1** and **Figure B. 2** show example relationships of the 200GAUI-8 and 400GAUI-16 C2C interfaces to the ISO/IEC OSI reference model, respectively. The 200GAUI-8 and 400GAUI-16 C2C interfaces provide electrical characteristics and associated compliance points, which can optionally be used when designing systems with electrical interconnect of approximately 25cm in length.



200GAUI-8 = 200 Gb/s EIGHT-LANE ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
200GXS = 200GMII EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure B. 1. *Example 200GAUI-8 chip-to-chip relationship to the ISO/IEC Open System Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model.*

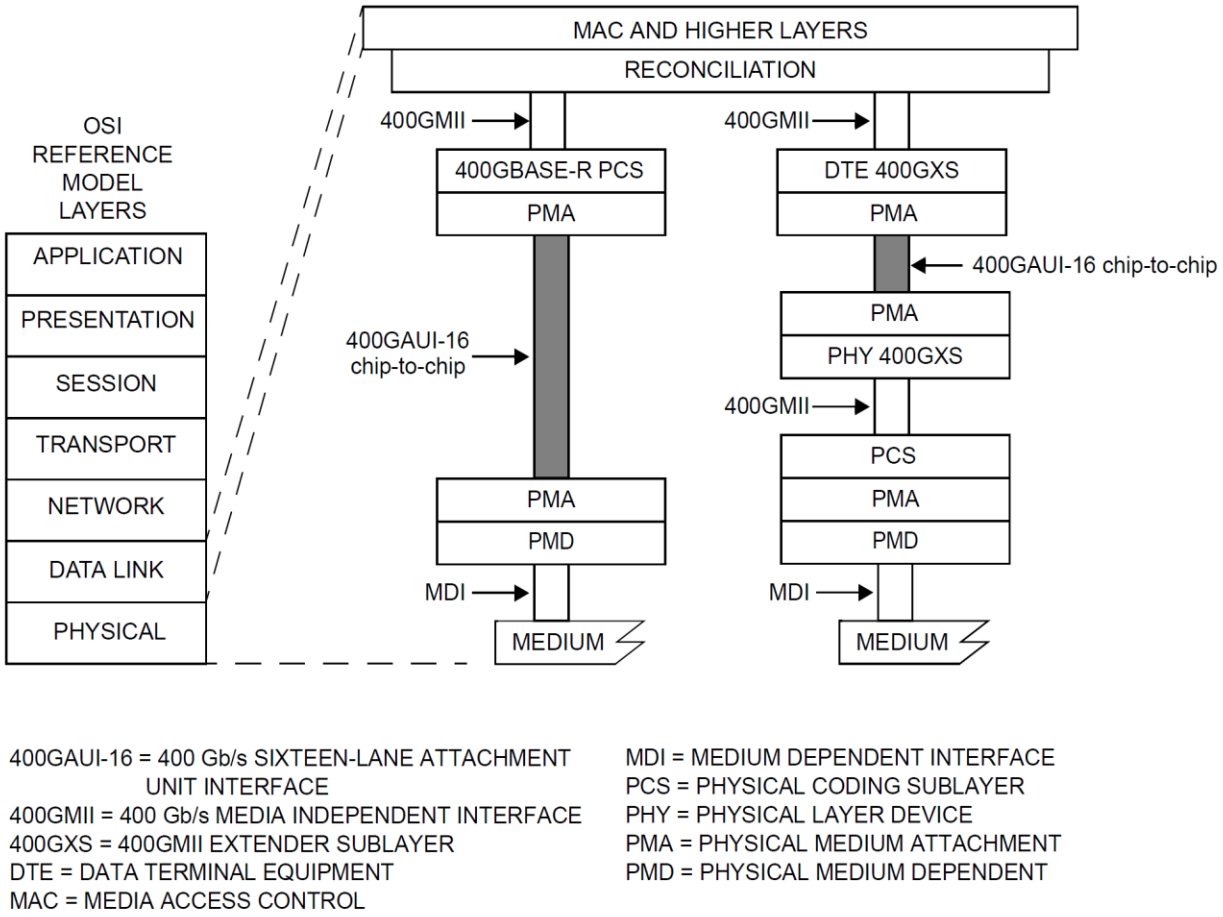


Figure B. 2. *Example 400GAUI-16 chip-to-chip relationship to the ISO/IEC Open System Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model.*

The 200GAUI-8 C2C link is described in terms of a 200GAUI-8 C2C transmitter, a 200GAUI-8 C2C channel, and a 200GAUI-8 C2C receiver. **Figure B. 4** depicts a typical 200GAUI-8 C2C application, and **Equation B- 1** as illustrated in **Figure B. 3**, summarizes the differential insertion loss budget associated with the chip-to-chip application. The 200GAUI-8 C2C interface comprises independent data paths in each direction. Each data path contains eight AC-coupled differential lanes using 2-level Non-Return-to-Zero (NRZ) (also known as PAM2) signaling, where the highest differential voltage level corresponds to the symbol one and the lowest level corresponds to the symbol zero. The low-frequency 3dB cutoff of the AC-coupling should be less than 100 kHz.

$$Insertion\_loss(f) \leq \begin{pmatrix} 1.083 + 2.543\sqrt{f} + 0.761f & 1.083 + 2.543\sqrt{f} + 0.761f \\ 1.083 + 2.543\sqrt{f} + 0.761f & 1.083 + 2.543\sqrt{f} + 0.761f \end{pmatrix}$$

Equation B- 1

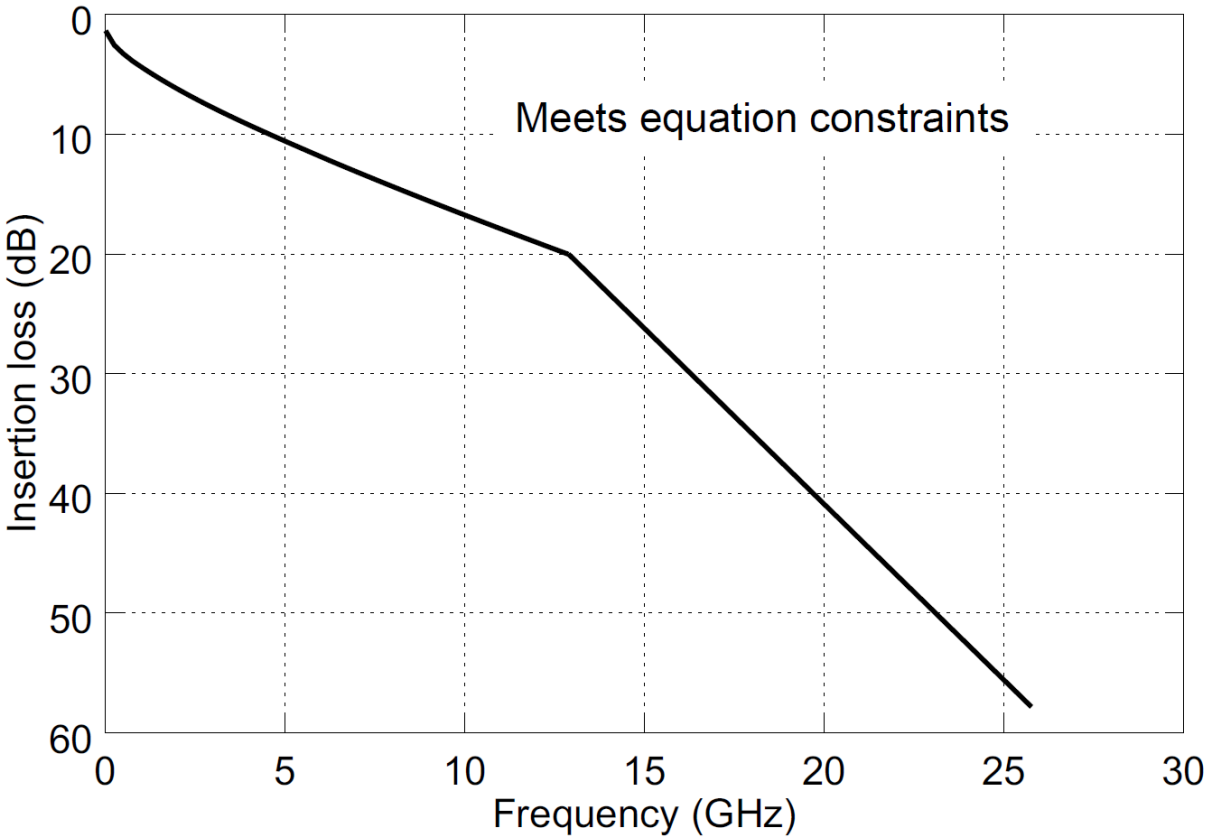


Figure B. 3. *CAUI-4 chip-to-chip channel insertion loss.*

The 400GAUI-16 C2C link is described in terms of a 400GAUI-16 C2C transmitter, a 400GAUI-16 C2C channel, and a 400GAUI-16 C2C receiver. **Figure B. 5** depicts a typical 400GAUI-16 C2C application, and Equation B-1, as illustrated in **Figure B. 3**, summarizes the informative differential insertion loss budget associated with the chip-to-chip application. The 400GAUI-16 C2C interface comprises independent data paths in each direction. Each data path contains sixteen AC-coupled differential lanes using 2-level NRZ (also known as PAM2) signaling, where the highest differential voltage level corresponds to the symbol one and the lowest level corresponds to the symbol zero. The low frequency 3dB cutoff of the Accoupling should be less than 100kHz.

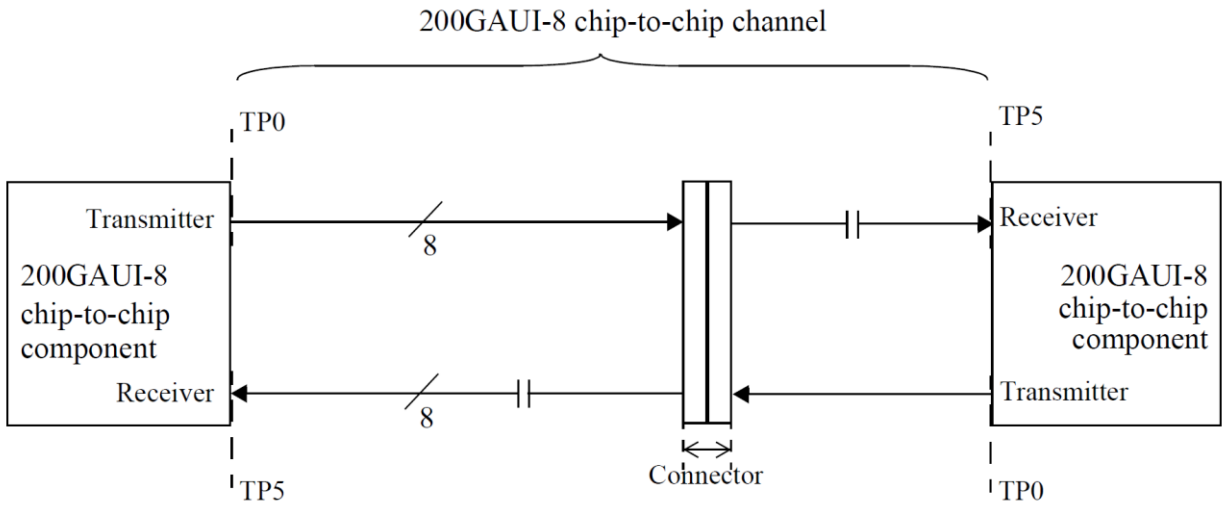


Figure B. 4. *Typical 200GAUI-8 chip-to-chip application.*

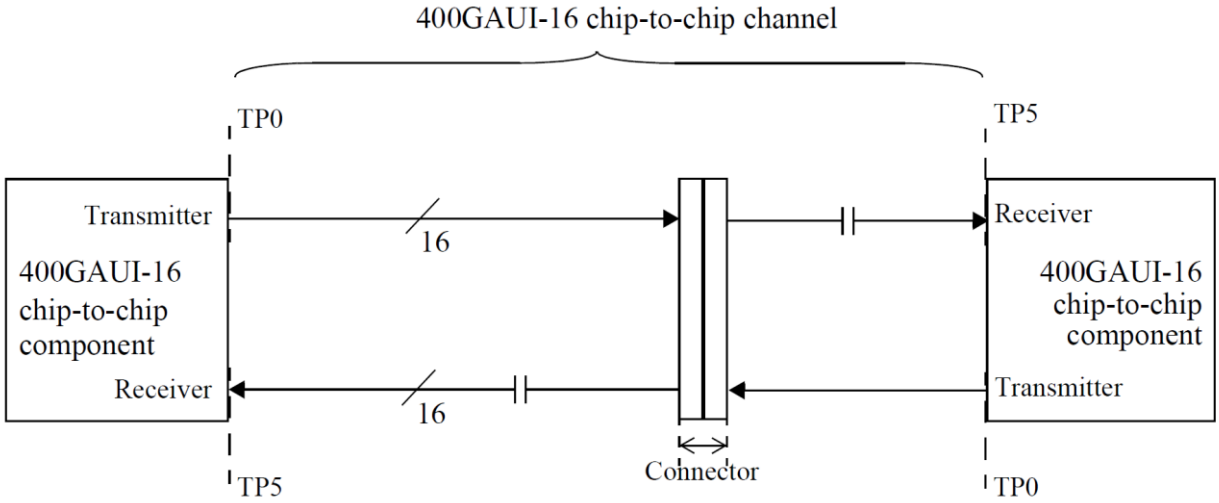


Figure B. 5. *Typical 400GAUI-16 chip-to-chip application.*

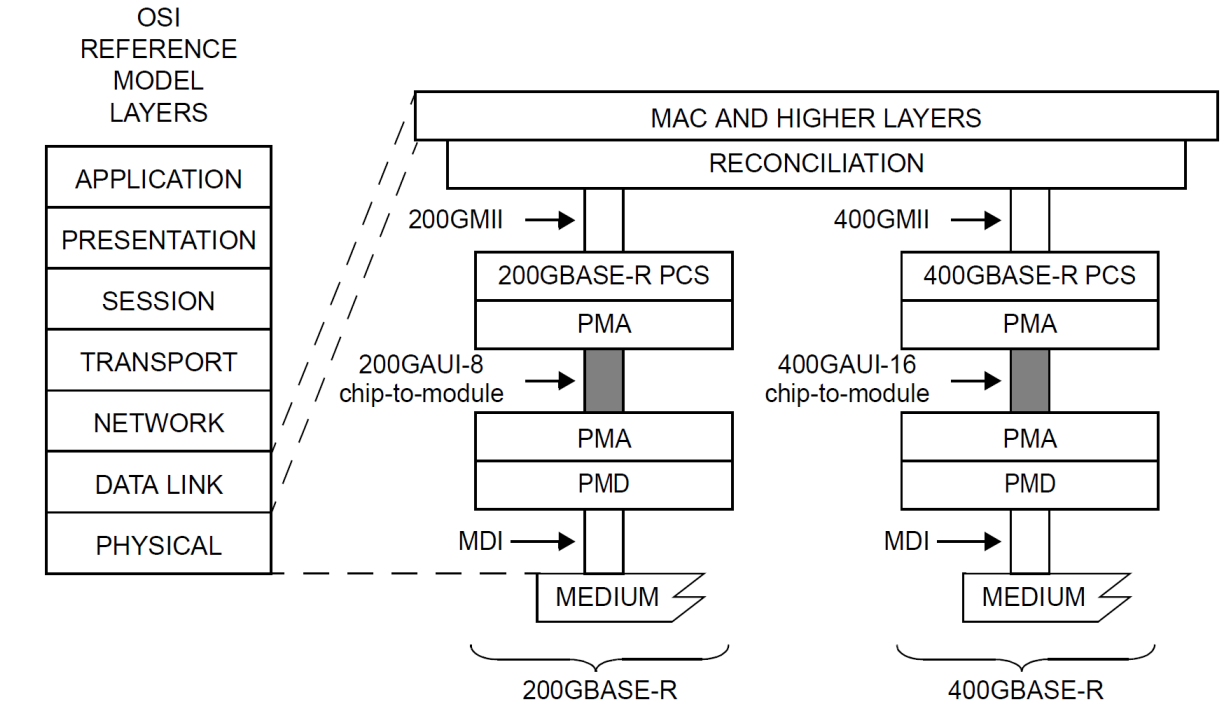
The nominal signaling rate for each lane is 26.5625GBd. The 200GAUI-8 C2C or 400GAUI-16 C2C transmitter on each end of the link is adjusted to an appropriate setting based on channel knowledge. If implemented, the transmitter equalization feedback mechanism may be used to identify an appropriate setting. The adaptive or adjustable receiver performs the remainder of the equalization.

The normative channel compliance is through 200GAUI-8 C2C or 400GAUI-16 C2C Channel Operating Margin (COM). Actual channel loss could be higher or lower than that

given by **Equation B- 1** due to the channel Insertion Loss Deviation (ILD), return loss, and crosstalk.

Annex C. Chip-to-module 200 Gb/s eight-lane Attachment Unit Interface (200GAUI-8 C2M) and 400 Gb/s sixteen-lane Attachment Unit Interface (400GAUI-16 C2M)

This annex defines the functional for the optional chip-to-module 200Gb/s eight-lane AUI (200GAUI-8 C2M) and 400Gb/s sixteen-lane AUI (400GAUI-16 C2M). **Figure C. 1** shows the relationship of the 200GAUI-8 and 400GAUI-16 C2M interfaces to the ISO/IEC OSI reference model. The C2M interface provides the electrical characteristics and the associated compliance points, which can optionally be used when designing systems with pluggable module interfaces.



200GAUI-8 = 200 Gb/s EIGHT-LANE ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI-16 = 400 Gb/s SIXTEEN-LANE ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure C. 1. *Example 200GAUI-8 and 400GAUI-16 chip-to-module relationship to the ISO/IEC Open System Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model.*

The sublayers (including the PCS and associated FEC) of each PHY that can optionally include a 200GAUI-8 C2M or 400GAUI-16 C2M are summarized in the tables in Section

3 and are specified in the corresponding PMD section. The positioning of the 200GAUI-8 C2M or 400GAUI-16 C2M relative to other sublayers is shown in Section **3.5** with further examples in 200Gb/s and 400Gb/s PMA sublayer partitioning examples.

The 200GAUI-8 C2M link is described in terms of a host 200GAUI-8 C2M component, a 200GAUI-8 C2M channel with associated insertion loss, and a module 200GAUI-8 C2M component. **Figure C. 2** depicts a typical 200GAUI-8 C2M application and summarizes the differential insertion loss budget associated with the C2M application. The supported insertion loss budget is characterized by **Equation B- 1**. The 200GAUI-8 C2M interface comprises independent data paths in each direction. Each data path contains eight differential lanes using 2-level NRZ (also known as PAM2) signaling, where the highest differential voltage level corresponds to the symbol one and the lowest level corresponds to the symbol zero. Each lane is AC-coupled within the module.

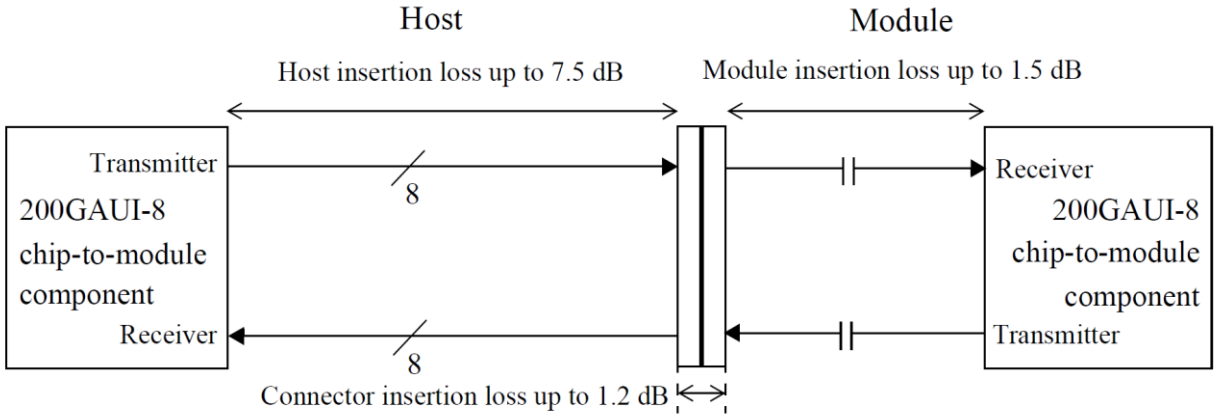


Figure C. 2. *200GAUI-8 chip-to-module insertion loss budget at 13.28GHz.*

The 400GAUI-16 C2M link is described in terms of a host 400GAUI-16 C2M component, a 400GAUI-16 C2M channel with associated insertion loss, and a module 400GAUI-16 C2M component. **Figure C. 3** depicts a typical 400GAUI-16 C2M application and summarizes the differential insertion loss budget associated with the chip-to-module application. The supported insertion loss budget is characterized by **Equation B- 1**. The 400GAUI-16 C2M interface comprises independent data paths in each direction. Each data path contains sixteen differential lanes using 2-level NRZ (also known as PAM2) signaling, where the highest differential voltage level corresponds to the symbol one and

the lowest level corresponds to the symbol zero. Each lane is AC-coupled within the module.

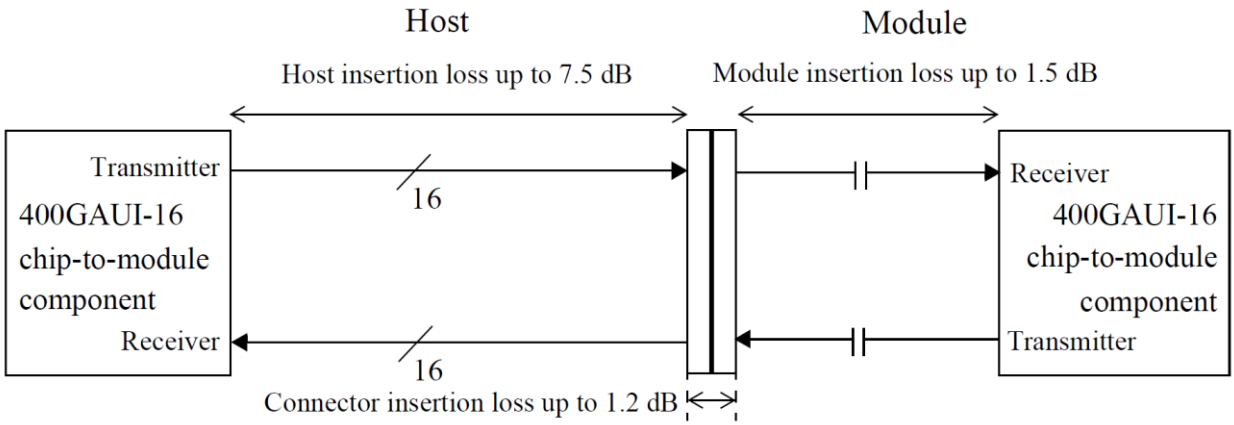
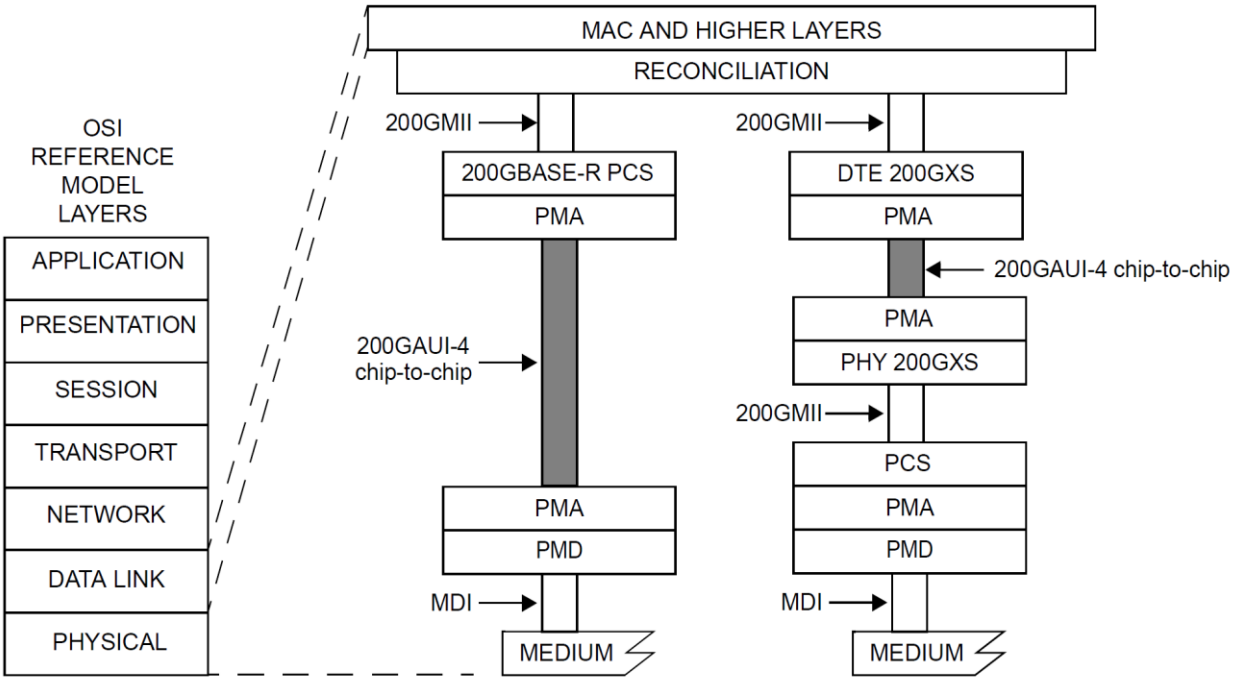


Figure C. 3. *400GAUI-16 chip-to-module insertion loss budget at 13.28GHz.*

The nominal signaling rate for each lane is 26.5625GBd. The 200GAUI-8 and 400GAUI-16 C2M interfaces are defined using a specification and test methodology that is similar to that used for CEI-28G-VSR defined in OIF-CEI-03.1. The bit error ratio (BER) shall be less than 10^{-6}.

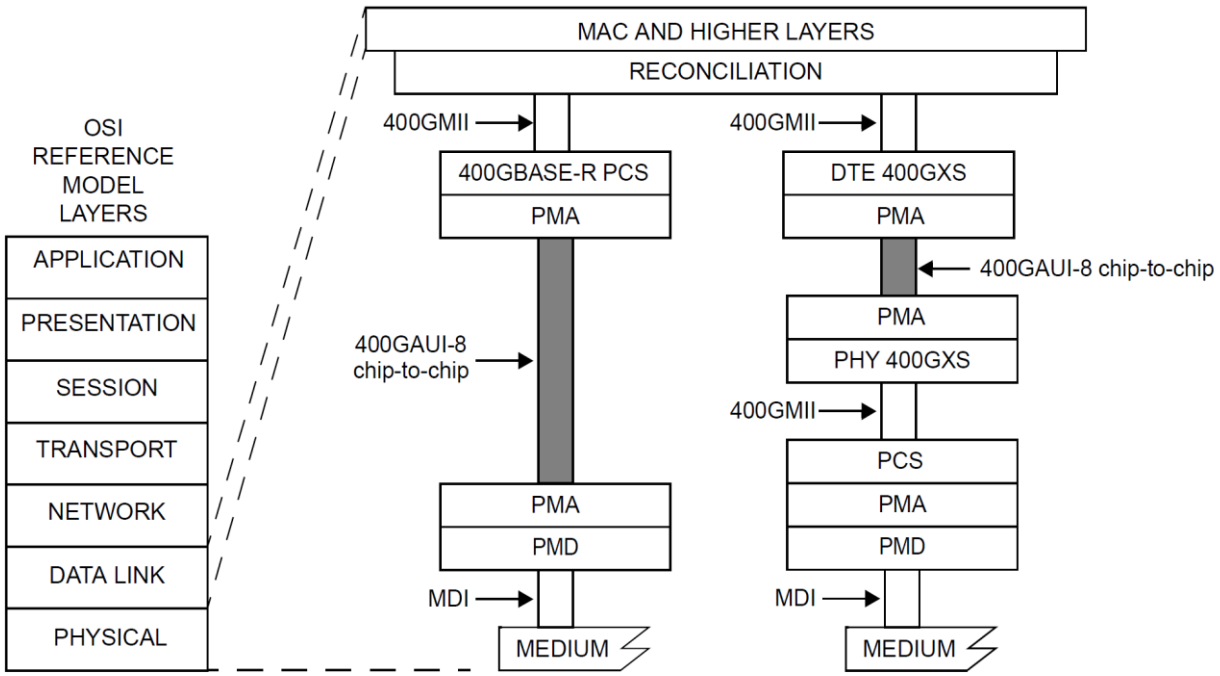
Annex D. Chip-to-chip 200 Gb/s four-lane Attachment Unit Interface (200GAUI-4 C2C) and 400 Gb/s eight-lane Attachment Unit Interface (400GAUI-8 C2C)

This annex defines the functional for the optional chip-to-chip 200Gb/s four-lane AUI (200GAUI-4 C2C) and 400Gb/s eight-lane AUI (400GAUI-8 C2C). **Figure D. 1** and **Figure D. 2** show example relationships of the 200GAUI-4 and 400GAUI-8 C2C interfaces to the ISO/IEC OSI reference model, respectively. The 200GAUI-4 and 400GAUI-8 C2C interfaces provide the electrical characteristics and the associated compliance points, which can optionally be used when designing systems with electrical interconnect of approximately 25cm in length.



200GAUI-4 = 200 Gb/s FOUR-LANE ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
200GXS = 200GMII EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure D. 1. *Example 200GAUI-4 chip-to-chip relationship to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model.*



400GAUI-8 = 400 Gb/s EIGHT-LANE ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
400GXS = 400GMII EXTENDER SUBLAYER
DTE = DATA TERMINAL EQUIPMENT
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure D. 2. *Example 400GAUI-8 chip-to-chip relationship to the ISO/IEC Open System Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model.*

The 200GAUI-4 C2C link is described in terms of a 200GAUI-4 C2C transmitter, a 200GAUI-4 C2C channel, and a 200GAUI-4 C2C receiver. **Figure D. 3** depicts a typical 200GAUI-4 C2C application, and **Equation D- 1** (illustrated in **Figure D. 5**) summarizes the informative differential insertion loss budget associated with the C2C application. The 200GAUI-4 C2C interface comprises independent data paths in each direction. Each data path contains four AC-coupled differential lanes using PAM4 signaling, where the highest differential level corresponds to the symbol three and the lowest level corresponds to the symbol zero. The low-frequency 3dB cutoff of the AC-coupling should be less than 100kHz.

The 400GAUI-8 C2C link is described in terms of a 400GAUI-8 C2C transmitter, a 400GAUI-8 C2C channel, and a 400GAUI-8 C2C receiver. **Figure D. 4** depicts a typical 400GAUI-8 C2C application, and **Equation D- 1** (illustrated in **Figure D. 5**) summarizes

the informative differential insertion loss budget associated with the chip-to-chip application. The 400GAUI-8 C2C interface comprises independent data paths in each direction. Each data path contains eight AC-coupled differential lanes using PAM4 signaling, where the highest differential level corresponds to the symbol three and the lowest level corresponds to the symbol zero. The low-frequency 3dB cutoff of the AC-coupling should be less than 100kHz.

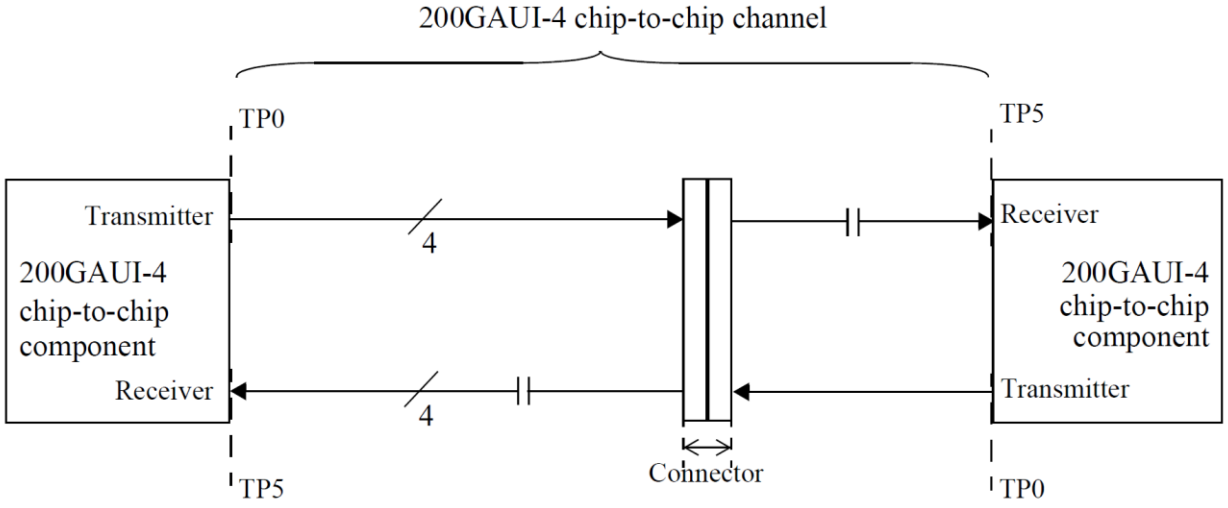


Figure D. 3. *Typical 200GAUI-4 chip-to-chip application.*

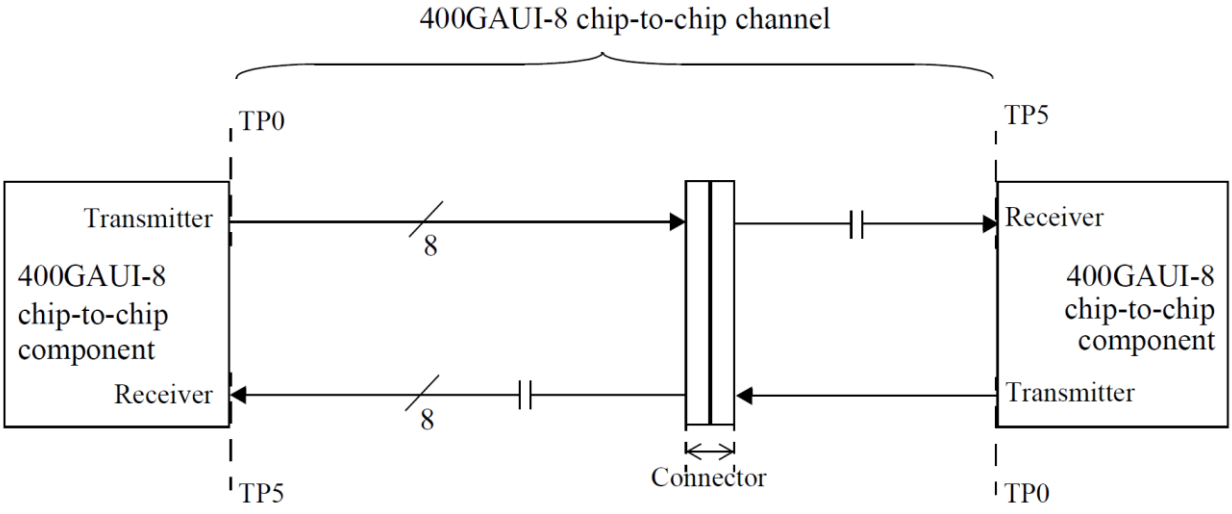


Figure D. 4. *Typical 400GAUI-8 chip-to-chip application.*

The nominal signaling rate for each lane is 26.5625GBd. The 200GAUI-4 C2C or 400GAUI-8 C2C transmitter on each end of the link is adjusted to an appropriate setting

based on channel knowledge. If implemented, the transmitter equalization feedback mechanism may be used to identify an appropriate setting. The adaptive or adjustable receiver performs the remainder of the equalization.

The channel is normatively defined using COM. Actual channel loss could be higher or lower than that given by **Equation D- 1** due to the channel ILD, return loss, and crosstalk. Note that for this equation the channel loss at the Nyquist frequency is lower than or equal to 20.457dB.

$$Insertion_{loss(f)} \leq 4.083 + 2.543\sqrt{f} + 0.761f \quad (d)$$

Equation D- 1

for $0.01 \leq f \leq 26.5625$

where:

f is the frequency in GHz,

$Insertion\_loss(f)$ is the informative 200GAUI-4 or 400GAUI-8 chip-to-chip insertion loss.

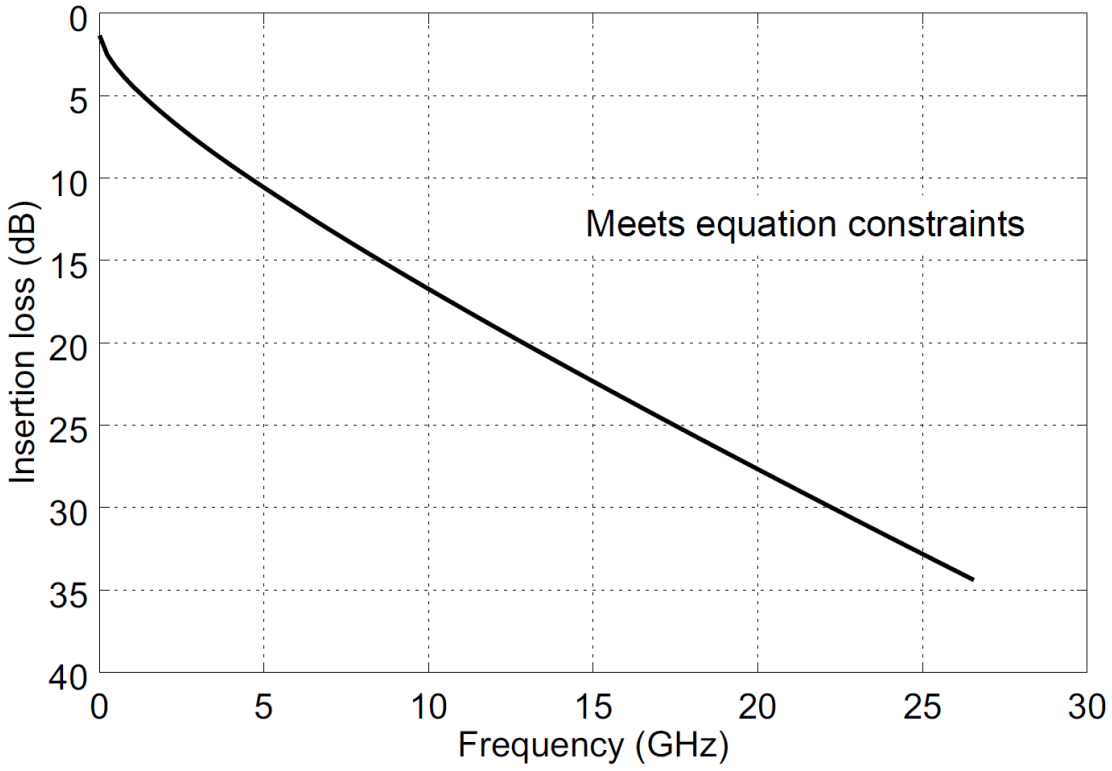
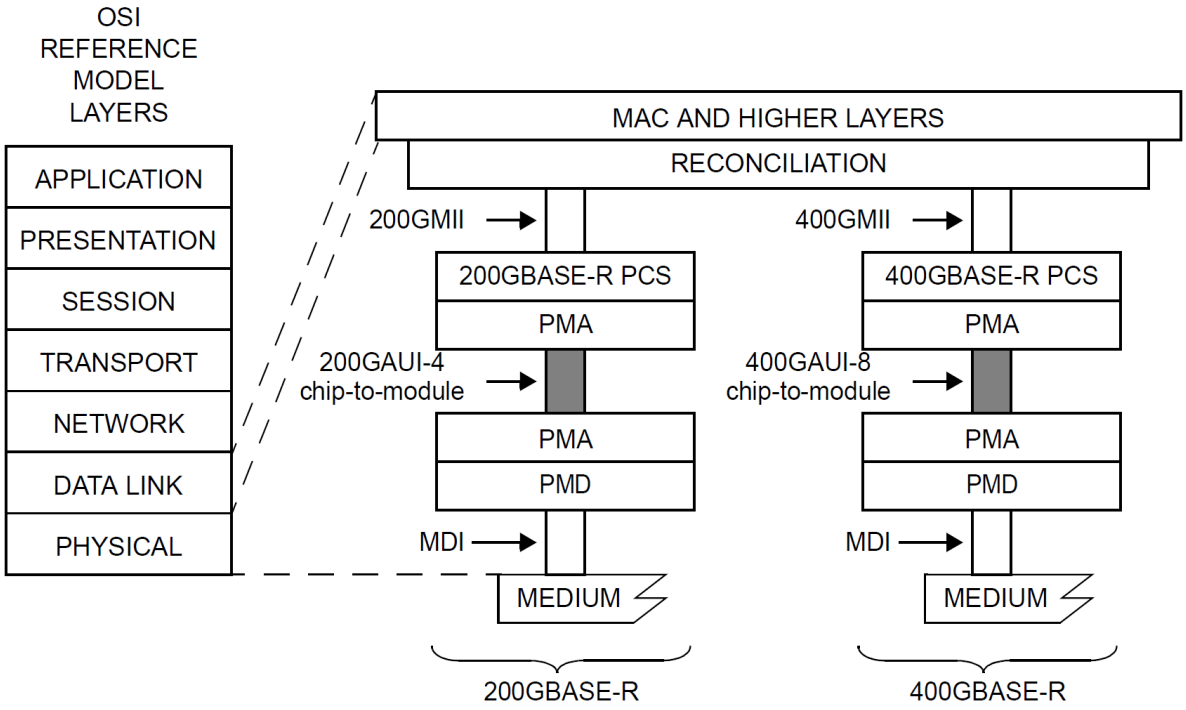


Figure D. 5. *200GAUI-4 or 400GAUI-8 chip-to-chip channel insertion loss.*

Annex E. Chip-to-module 200 Gb/s four-lane Attachment Unit Interface (200GAUI-4 C2M) and 400 Gb/s eight-lane Attachment Unit Interface (400GAUI-8 C2M)

This annex defines the functional for the optional chip-to-module 200 Gb/s four-lane AUI (200GAUI-4 C2M) and 400 Gb/s eight-lane AUI (400GAUI-8 C2M). **Figure E. 1** shows the relationship of the 200GAUI-4 and 400GAUI-8 C2M interfaces to the ISO/IEC OSI reference model. The C2M interface provides the electrical characteristics and the associated compliance points, which can optionally be used when designing systems with pluggable module interfaces.



200GAUI-4 = 200 Gb/s FOUR-LANE ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI-8 = 400 Gb/s EIGHT-LANE ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure E. 1. *Example 200GAUI-4 and 400GAUI-8 chip-to-module relationship to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model.*

The sublayers (including the PCS and associated FEC) for each PHY that can optionally include a 200GAUI-4 C2M or 400GAUI-8 C2M are summarized in the tables in Section **3.1.3** and are specified in the corresponding PMD section. The positioning of the 200GAUI-4 C2M or 400GAUI-8 C2M relative to other sublayers is shown in Section **3.5** with further examples in **Annex A**.

The 200GAUI-4 C2M link is described in terms of a host 200GAUI-4 C2M component, a 200GAUI-4 C2M channel with associated insertion loss, and a module 200GAUI-4 C2M component. **Figure E. 2** depicts a typical 200GAUI-4 C2M application and summarizes the differential insertion loss budget associated with the chip-to-module application. The supported insertion loss budget is characterized by **Equation E- 1** and illustrated in **Figure E. 4**. The 200GAUI-4 C2M interface comprises independent data paths in each direction. Each data path contains four differential lanes using PAM4 signaling, where the highest differential level corresponds to the symbol three and the lowest level corresponds to the symbol zero. Each lane is AC-coupled within the module.

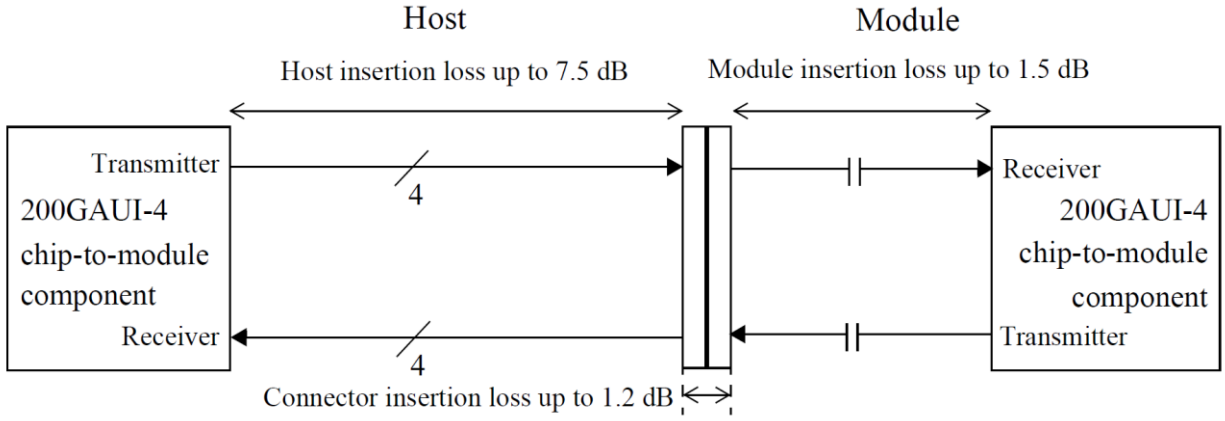


Figure E. 2. *200GAUI-4 chip-to-module insertion loss budget at 13.28GHz.*

The 400GAUI-8 C2M link is described in terms of a host 400GAUI-8 C2M component, a 400GAUI-8 C2M channel with associated insertion loss, and a module 400GAUI-8 C2M component. **Figure E. 3** depicts a typical 400GAUI-8 C2M application and summarizes the differential insertion loss budget associated with the chip-to-module application. The recommended insertion loss budget is characterized by **Equation E- 1** and illustrated in **Figure E. 4**. The 400GAUI-8 C2M interface comprises independent data paths in each direction. Each data path contains eight differential lanes using PAM4 signaling, where

the highest differential level corresponds to the symbol three and the lowest level corresponds to the symbol zero. Each lane is AC-coupled within the module.

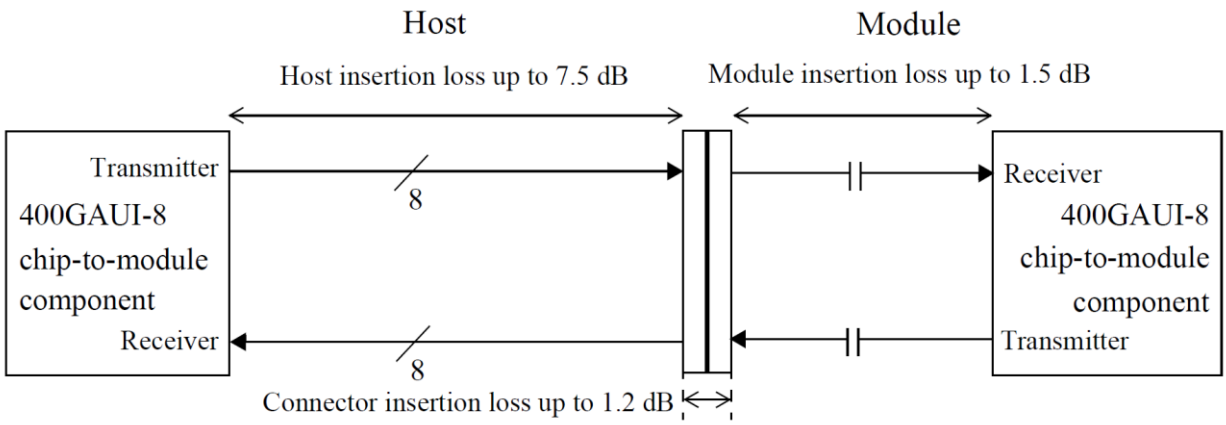


Figure E. 3. *400GAUI-8 chip-to-module insertion loss budget at 13.28GHz.*

The nominal signaling rate for each lane is 26.5625GBd. The C2M interface is defined using a specification and test methodology that is similar to that used for CEI-56G-VSR-PAM defined in OIF-CEI-04.0 [B55a].

$$Insertion_{loss(f)} \leq \begin{pmatrix} 0.05 + 1.8\sqrt{f} + 0.270f & 0.01 \leq f \leq 13.28 \\ -4.0096 + 1.07f & 13.28 < f \leq 26.5625 \end{pmatrix} (dB)$$

Equation E- 1

where:

f is the frequency in GHz,

$Insertion\_loss(f)$ is the informative 200GAUI-4 or 400GAUI-8 chip-to-module insertion loss.

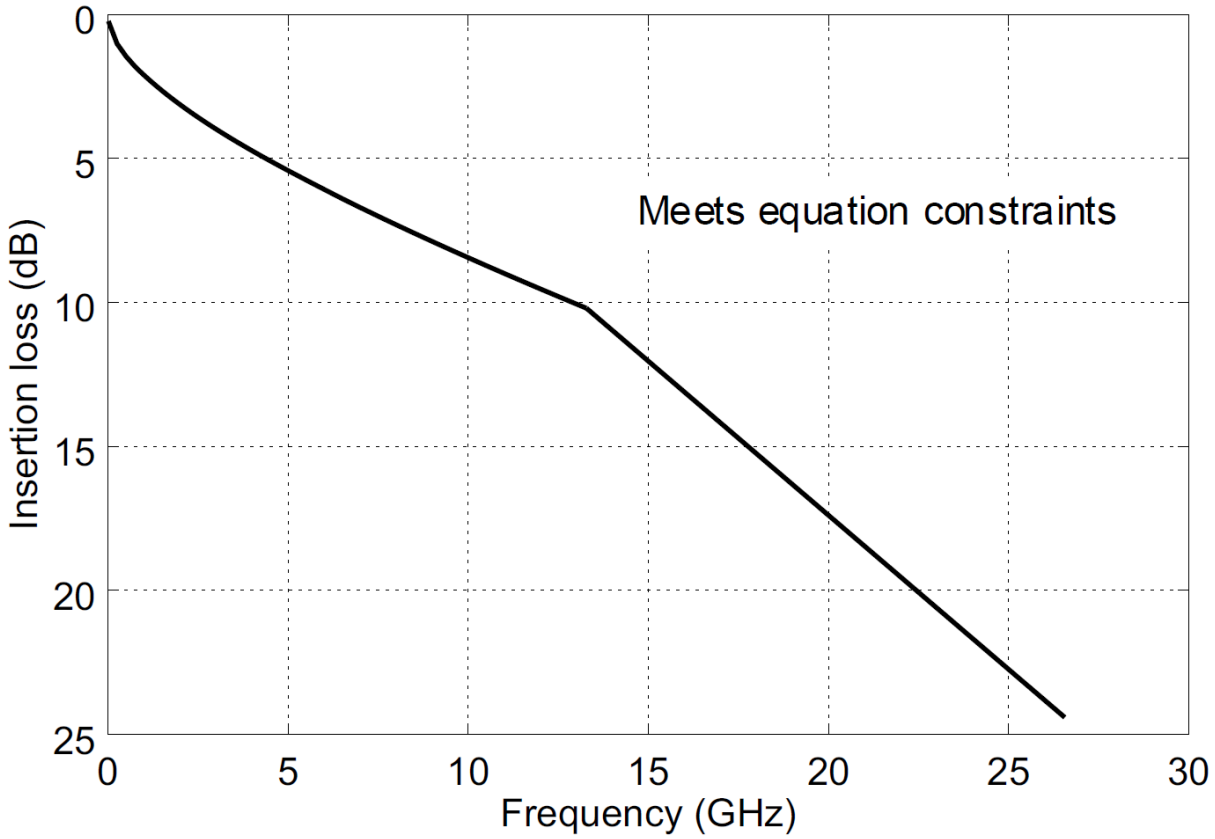


Figure E. 4. *Recommended 200GAUI-4 or 400GAUI-8 chip-to-module channel insertion loss.*

The BER when processed according to Section **3.5** shall be less than 10^{-5}.

Annex F. Energy-Efficient Ethernet (EEE)

The optional EEE capability combines the IEEE 802.3 MAC Sublayer with a family of PHYs defined to support operation in the LPI mode. When the LPI mode is enabled, systems on both sides of the link can save power during periods of low link utilization.

EEE also provides a protocol to coordinate transitions to or from a lower level of power consumption and does this without changing the link status and without dropping or corrupting frames. The transition time into and out of the lower level of power consumption is kept small enough to be transparent to upper layer protocols and applications.

EEE supports operation over twisted-pair cabling systems, twinaxial cable, electrical backplanes, optical fiber, the CAUI-10 or CAUI-4 for 100 Gb/s PHYs, the 200GAUI-n and 200GXS for 200 Gb/s PHYs, and the 400GAUI-n and 400GXS for 400 Gb/s PHYs.

EEE also specifies means to exchange capabilities between link partners to determine whether EEE is supported and to select the best set of parameters common to both devices.

F.1 LPI Signaling

LPI signaling allows the LPI Client to indicate to the PHY, and to the link partner, that a break in the data stream is expected, and the LPI Client can use this information to enter power-saving modes that require additional time to resume normal operation. LPI signaling also informs the LPI Client when the link partner has sent such an indication.

The definition of LPI signaling assumes the use of the MAC for simplified full duplex operation (with carrier sense deferral). This provides full duplex operation but uses the carrier sense signal to defer transmission when the PHY is in the LPI mode.

The LPI Client connects to the RS service interface. LPI signaling between the RS and PCS is performed by LPI encoding on the MII. The transmit PCS encodes LPI symbols, which are decoded by the link partner receive PCS. The receive and transmit PCS also generate service interface signals, which are passed down to the lower PHY sublayers and indicate when receive and transmit PHY functions may be powered down.

The EEE request signals from the PCS control transitions between quiescent and normal operation. The PCS sublayer also requests transmit alert operation to assist the partner device PMD to detect the end of the quiescent state. Additionally, these PCS

types generate the *RX_LPI_ACTIVE* signal, which indicates to BASE-R FEC that it can use rapid block lock because the link partner PCS has bypassed scrambling.

Coding also allows LPI transmit quiet and alert requests from the PCS to be signaled over the CAUI-n interfaces. The CAUI-n receive interfaces infer the quiet and alert requests from the data received over the interface and use that to recreate the transmit or receive direction signaling.

The receive PCS checks that the link cycles out of the quiescent state at the correct time and that the received signals return to their expected state within the required time. The ENERGY_DETECT indicate signal is passed up from the PMA to the PCS to allow the PCS to monitor the waking process.

F.1.1 Reconciliation sublayer service interfaces

Figure F. 1 depicts the LPI Client and the RS interlayer service interfaces.

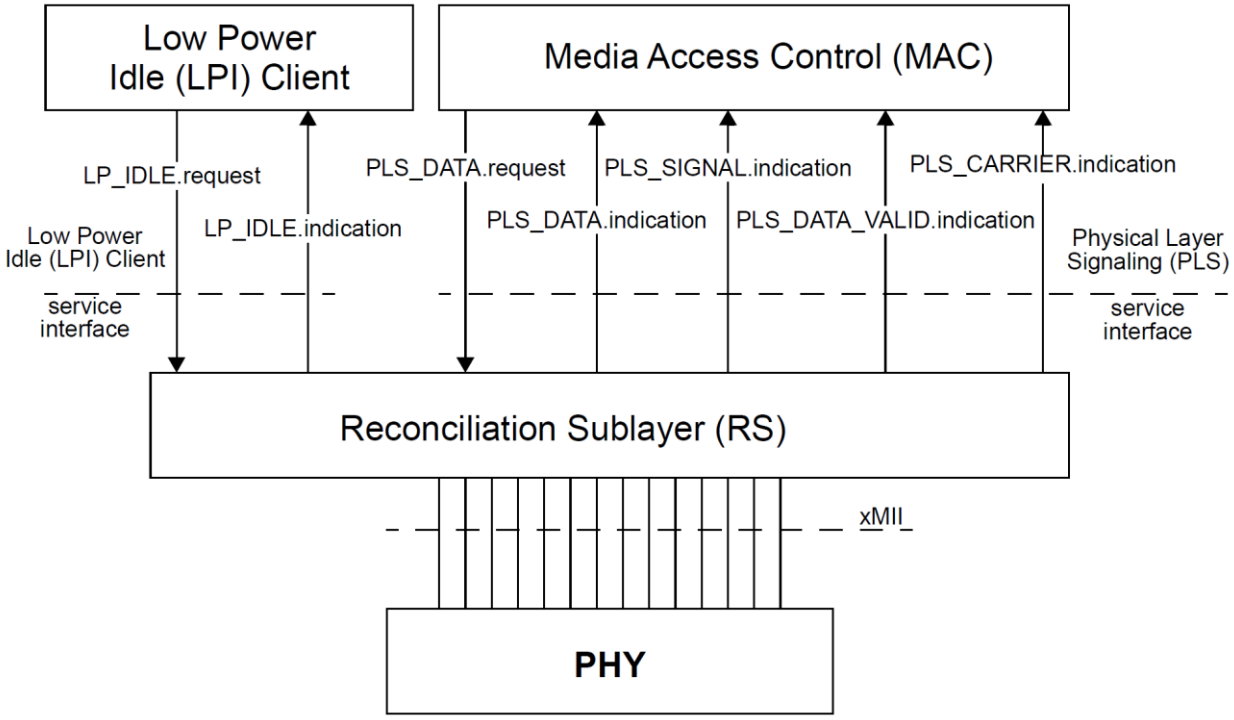


Figure F. 1. *LPI Client and RS interlayer service interfaces.*

F.1.2 Reconciliation sublayer operation

LPI assert and detect functions are contained in the Reconciliation Sublayer as shown in **Figure F. 2**. The xMII in this diagram represents any of the family of medium independent interfaces supported by EEE.

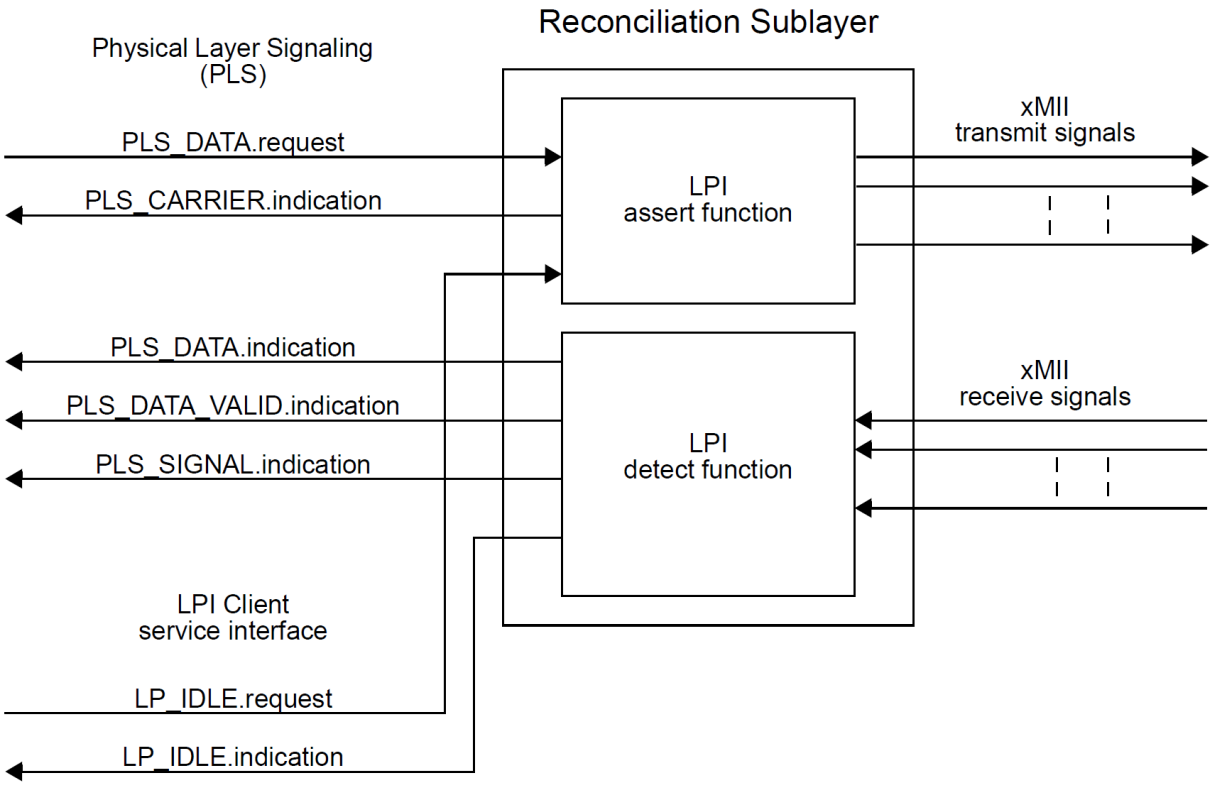


Figure F. 2. *RS LPI assert and detect functions.*

The following provides an overview of RS LPI operation.

F.1.3 PHY LPI transmit operation

When the “Assert LPI” encoding on the xMII is detected, the PHY signals sleep to its link partner, indicating that the local transmitter is entering the LPI mode.

The EEE capability in most PHYs requires the local PHY transmitter to be quiet after sleep is signaled.

The transmit function of the local PHY is periodically enabled to transmit refresh signals, these signals are used by the link partner to update the adaptive filters and the timing of the circuits, in order to maintain link integrity.

This cycle of quiet refresh continues until the reception of the normal interframe encoding on the xMII. The transmit function in the PHY communicates this state to the link partner by sending a wake signal for a predefined period of time. The PHY then enters the normal operating state.

Figure F. 3 shows the general principles of the EEE-capable transmitter operation.

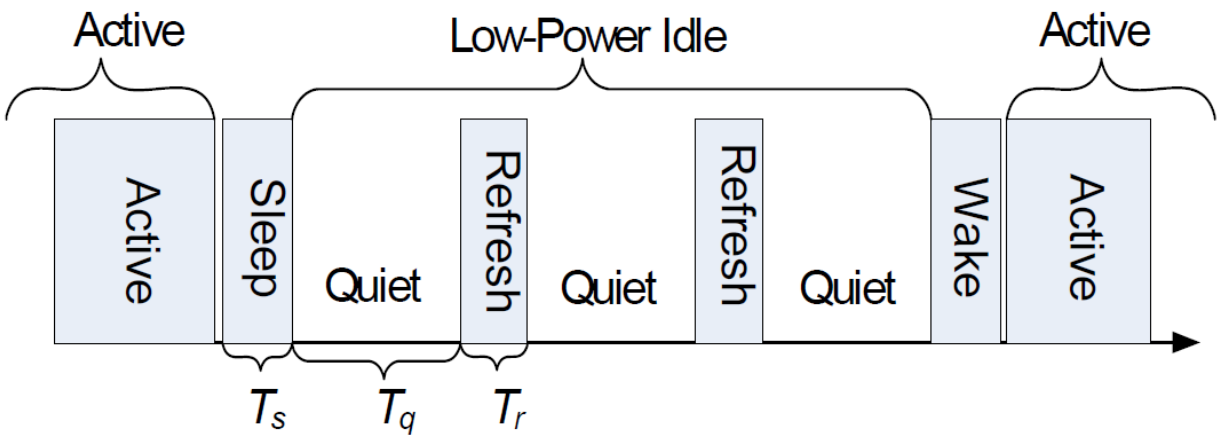


Figure F. 3. *Overview of EEE LPI operation.*

During the transition to or from the LPI mode, no data frames are lost or corrupted.

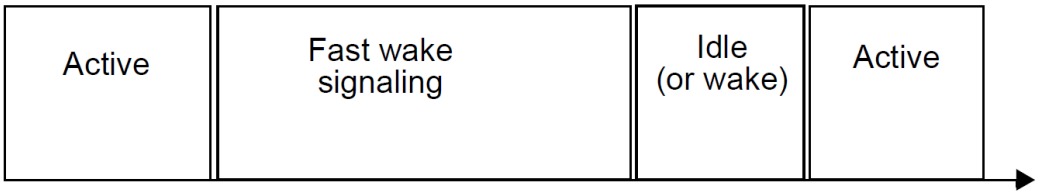
Except for BASE-T, the PHYs with an operating speed of 25Gb/s or greater, which implement the optional EEE capability, support two modes of LPI operation: deep sleep and fast wake. The deep sleep operation refers to the mode for which the transmitter ceases transmission during LPI (as shown in **Figure F. 3**), The deep sleep operation is equivalent to the mechanism defined for PHYs with an operating speed of 10Gb/s or below. Also, the deep sleep operation is optional for PHYs with an operating speed of 25Gb/s or greater, which implement EEE with the exception of the PHYs noted in **Table F. 1**, which do not support the deep sleep operation. The fast wake operation refers to the mode for which the transmitter continues to transmit signals during LPI, so that the receiver can resume operation with a shorter wake time (as shown in **Figure F. 4**). For transmit, other than the PCS encoding LPI, there is no difference between the fast wake and the normal operation.

PHY or interface type	Section
100GBASE-SR4 [a]	**2.4, 2.5, 2.8, 2.9**
100GBASE-SR10 [a]	**2.4, 2.5, 2.6**
100GBASE-LR4 [a]	**2.4, 2.5, 2.7**
100GBASE-ER4 [a]	**2.4, 2.5, 2.7**
200GBASE-DR4 [a]	**3.4, 3.5, 3.6**
200GBASE-FR4 [a]	**3.4, 3.5, 3.7**
200GBASE-LR4 [a]	**3.4, 3.5, 3.7**

400GBASE-SR16 [a]	**3.4, 3.5, 3.8**
400GBASE-DR4 [a]	**3.4, 3.5, 3.9**
400GBASE-FR8 [a]	**3.4, 3.5, 3.7**
400GBASE-LR8 [a]	**3.4, 3.5, 3.7**
[a] The deep sleep mode of EEE may not supported for this PHY.	

Table F. 1. *Sections associated with each PHY or interface type*

The PHY signaling continues with higher layer functions suspended during the fast wake signaling.



NOTE—Fast wake signaling continually indicates LPI in a normally constituted data stream.

Figure F. 4. *Overview of fast wake operation.*

F.1.4 PHY LPI receive operation

In the receive direction, the LPI mode is triggered by the reception of a sleep signal from the link partner, which indicates that the link partner is entering the LPI mode. After sending the sleep signal, the link partner ceases transmission, if it is not in fast wake mode. When the receiver detects the sleep signal, the local PHY assigns the “Assert LPI” on the xMII and the local receiver disables some functionality to reduce the power consumption.

If the link partner is not in fast wake mode, it periodically transmits refresh signals that are used by the local PHY to update the adaptive coefficients and the timing of the circuits. This quiet-refresh cycle continues until the link partner initiates transition back to the normal mode. This is implemented by transmitting the wake signal for a predetermined period of time controlled by the LPI assert function in the RS. This also allows the local receiver to be prepared for normal operation and to prepare the transition of the “Assert LPI” encoding to the normal interframe encoding on the xMII.

Annex G. MAC Control PAUSE operation

The PAUSE operation is used to restrict the transmission of data frames for a specified period of time. A MAC Control client wishing to inhibit transmission of data frames from another station on the network generates a *MA_CONTROL.request* primitive specifying the following:

a. The globally assigned 48-bit multicast address 01-80-C2-00-00-01.
b. The PAUSE opcode.
c. A *request_operand* indicating the length of time for which it wishes to inhibit data frame transmission.

The PAUSE operation cannot be used for the restriction of the MAC Control frames transmission.

The PAUSE frames shall only be sent by the DTEs, which are configured to the full duplex mode of operation.

The 48-bit multicast address 01-80-C2-00-00-01, which is globally assigned, has been reserved for use in MAC Control PAUSE frames. These MAC Control frames restrict the transmission of the data frames from a DTE in a full duplex mode IEEE 802.3 LAN. To allow the generic full duplex flow control, the stations implementing the PAUSE operation shall instruct the MAC (e.g., through the layer management) to enable the reception of frames with the destination address equal to this multicast address.

NOTE: The use of a well-known multicast address relieves the MAC Control sublayer and its client from having to know, and maintain knowledge of, the individual 48-bit address of the other DTE in a full duplex environment.

When MAC Control Priority-based Flow Control (PFC) operation has been enabled, MAC Control PAUSE operation shall be disabled.

The PAUSE opcode takes the following *request_operand*:

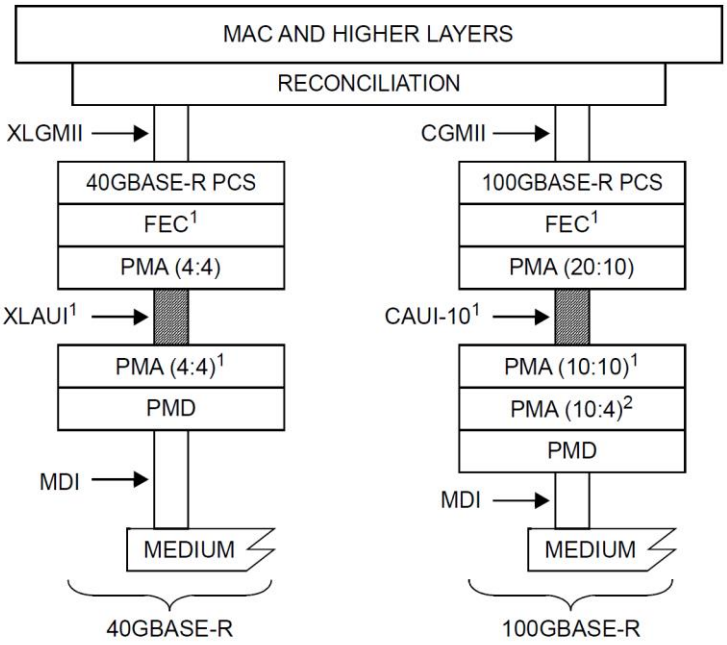
pause_time

A 2-octet, unsigned integer containing the length of time for which the receiving station is requested to restrict the data frame transmission. The most-significant octet is transmitted first, and the least-significant octet is transmitted second. The *pause_time* is measured in *pause_quanta* units, which are equal to 512-bit times

of the implementation. The range of possible *pause_time* is between 0 and 65535 *pause_quanta*.

Annex H. 100 Gb/s ten-lane Attachment Unit Interface (CAUI-10)

This annex defines the functional and electrical characteristics for the 100Gb/s ten-lane Attachment Unit Interface (CAUI-10). **Figure H. 1** shows the relationships of the CGMII, PMA, CAUI-10, and PMD for 100Gb/s.



CAUI-10 = 100 Gb/s TEN-LANE ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLAUI = 40 Gb/s ATTACHMENT UNIT INTERFACE
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

NOTE 1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE
NOTE 2—CONDITIONAL BASED ON PMD TYPE

Figure H. 1. *Example relationship of XLAUI and CAUI-10 to IEEE 802.3 Ethernet model.*

The purpose of the CAUI-10 is to provide a flexible chip-to-chip and chip-to-module interconnect for 100Gb/s components. This annex provides compliance requirements for CAUI-10 transmitters and receivers, while **Annex I** specifies the electrical requirements for the C2M interconnection.

An example application of CAUI-10 includes providing a physical connection between a ten-lane 100Gb/s PMA and a 10:4 PMA mapping element.

The CAUI-10 interface has the following characteristics:

a. Independent transmit and receive data paths.
b. Differential AC-coupled signaling with low voltage swing.
c. Self-timed interface.
d. Shared technology with other 40Gb/s or 100Gb/s interfaces.
e. Utilization of 64B/66B coding.

The following is a list of the major concepts of CAUI-10:

a. The CAUI-10 interface can be inserted between PMA layers in the IEEE 802.3 Ethernet model to transparently enable chip-to-chip communication.
b. The CAUI-10 is organized into ten lanes.
c. The CAUI-10 interface is a parallel electrical interface with each lane running at a nominal rate of 10.3125Gb/s.

The CAUI-10 interface supports the 100Gb/s data rate. For 40Gb/s applications, the data stream shall be presented in four lanes as described in Section **2.5**. For 100Gb/s applications, the data stream shall be presented in ten lanes as described in Section **2.5**. The data is 64B/66B coded. The nominal signaling rate for each lane in 100Gb/s applications shall be 10.3125Gb/s.

H.1. CAUI-10 link block diagram

CAUI-10 link is illustrated in **Figure H. 2**. CAUI-10 channel is defined from the transmit pad to the receive pad including any AC-coupling in the path.

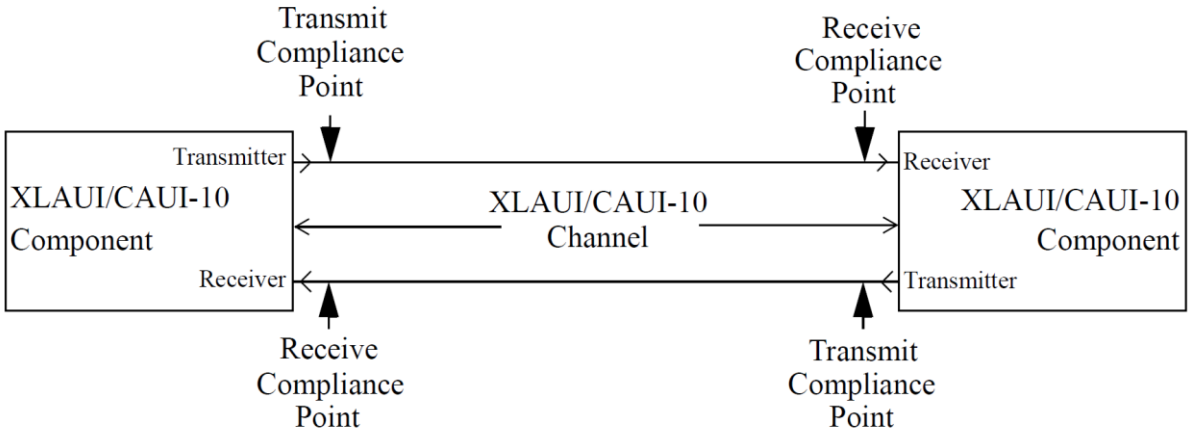


Figure H. 2. *Definition of transmit and receive compliance points.*

H.1.1. Transmitter compliance points

The reference differential insertion loss, expressed in decibels, between the transmitter and the transmit compliance point is defined in **Equation H- 1** and illustrated in **Figure H. 3**. The effects of differences between the actual insertion loss and the reference insertion loss are to be accounted in the measurements.

$$Insertion\_loss(f) = -0.00086 + 0.2286\sqrt{f} + 0.08386f \quad dB \qquad 0.01 \le f \le 11.1$$

Equation H- 1

where

Insertion_loss(f) is the differential insertion loss at frequency *f*

f is the frequency in GHz

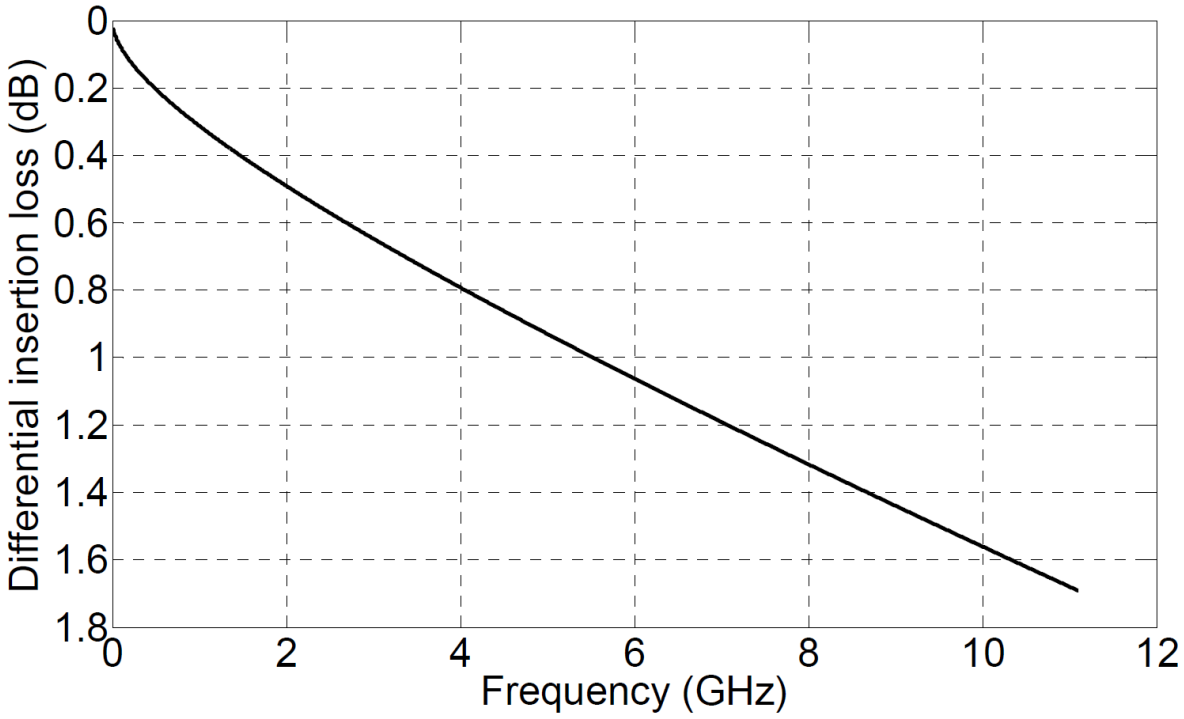


Figure H. 3. *Insertion loss between transmitter and transmit compliance point.*

H.1.2. Receiver compliance points

The reference differential insertion loss, expressed in decibels, between the receiver compliance point and the receiver is defined in **Equation H- 2** and illustrated in **Figure H. 4**. The effects of differences between the actual insertion loss and the reference insertion loss are to be accounted in the measurements.

$$Insertion\_loss(f) = -0.00086 + 0.2286\sqrt{f} + 0.08386f \quad dB \qquad 0.01 \le f \le 11.1$$

Equation H- 2

where

Insertion_loss(f) is the differential insertion loss at frequency *f*

f is the frequency in GHz

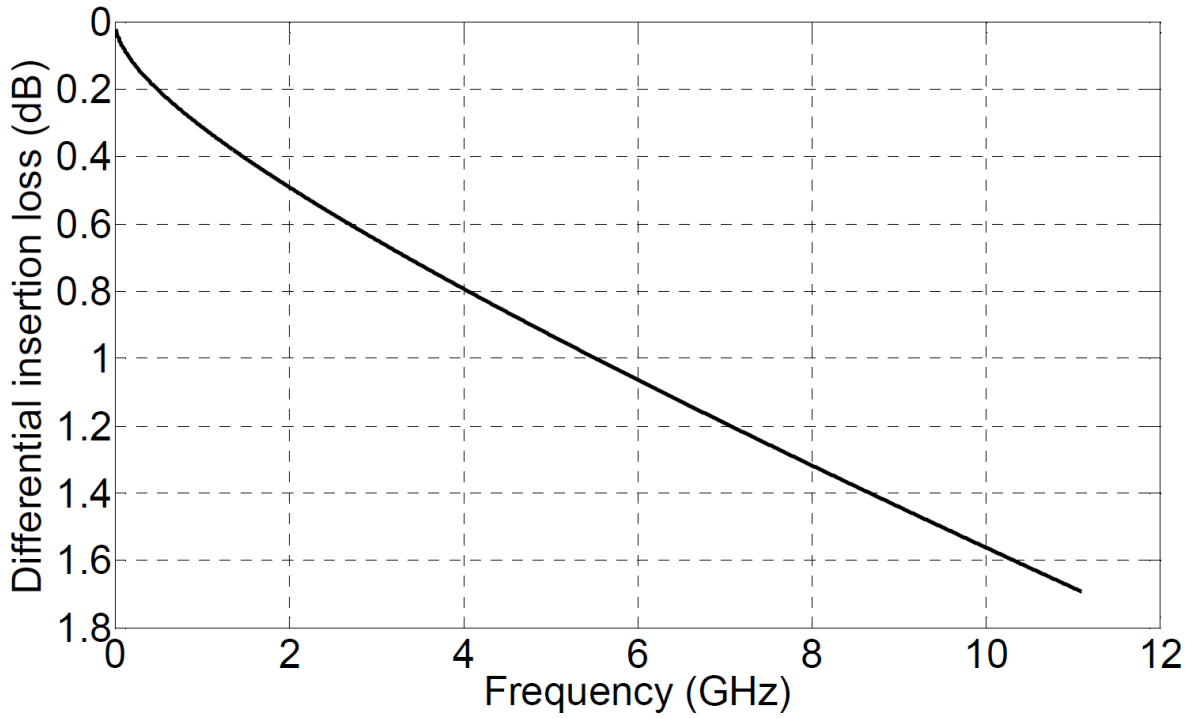


Figure H. 4. *Insertion loss between receive compliance point and receiver.*

Annex I. Chip-to-module 100Gb/s ten-lane Attachment Unit Interface (CAUI-10)

This annex defines the functional and electrical characteristics for the 100Gb/s ten-lane Attachment Unit Interface (CAUI-10). The purpose of this annex is to provide electrical characteristics and associated compliance points for pluggable module applications that use the CAUI-10 interface and shall use the same number of lanes and signaling rate defined in **Annex H**. **Figure I. 3** and **Table I. 1** summarize an example differential insertion loss budget associated with the C2M application. The insertion loss excluding a 0.5dB connector loss at 5.15625GHz, is linearly scaled to 7.9dB loss at 5.15625GHz for the host XLAUI/CAUI-10 component, and 2.1dB loss at 5.15625GHz for the module as per **Table I. 1.** Chip-to-module loss budget. and **Equation I- 1** for the host and **Equation I- 2** for the module. Tradeoffs can be made between the host PCB insertion loss and the connector loss. **Equation I- 1** is illustrated in **Figure I. 1** and **Equation I- 2** is illustrated in **Figure I. 2**.

$$Insertion\_loss(f) \leq \begin{Bmatrix} 0.111 + 1.046\sqrt{f} + 1.05f & 0.01 \leq f < 7 \\ -11.82 + 3.15f & 7 \leq f \leq 11.1 \end{Bmatrix} \ (dB)$$

Equation I- 1

where

Insertion_loss(f) is the differential insertion loss at frequency *f*

f is the frequency in GHz

$$Insertion\_loss(f) \leq \begin{Bmatrix} 0.03 + 0.278\sqrt{f} + 0.28f & 0.01 \leq f < 7 \\ -3.155 + 0.84f & 7 \leq f \leq 11.1 \end{Bmatrix} \ (dB)$$

Equation I- 2

where

Insertion_loss(f) is the differential insertion loss at frequency *f*

f is the frequency in GHz

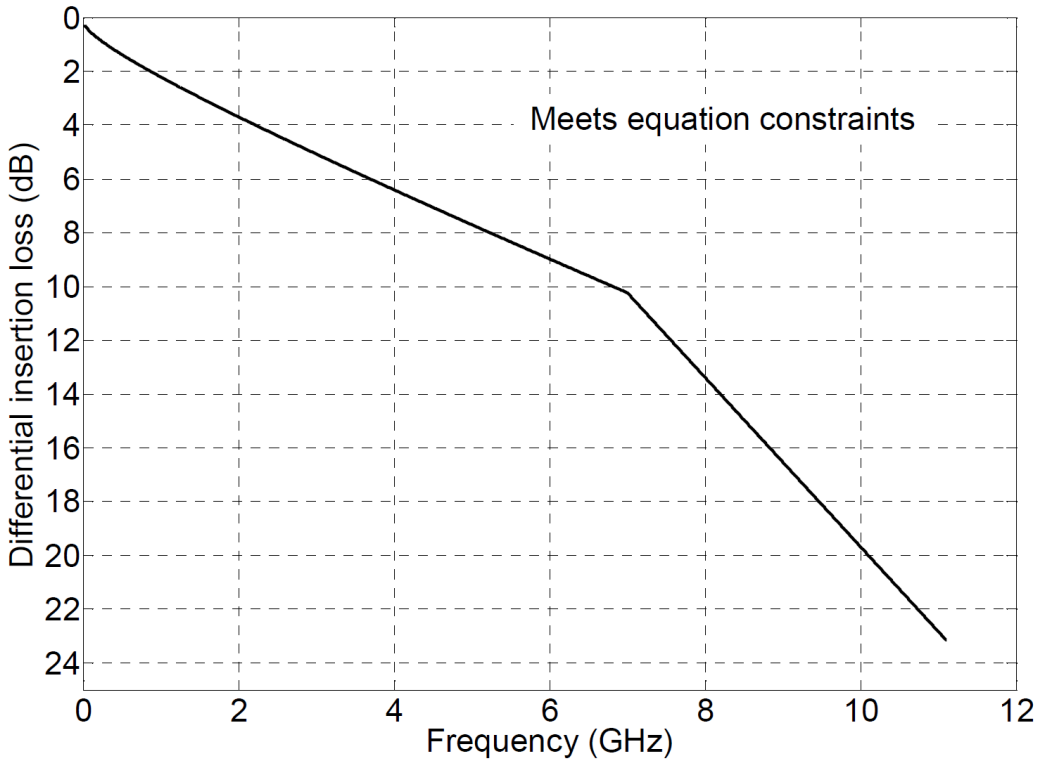


Figure I. 1. *Host insertion loss.*

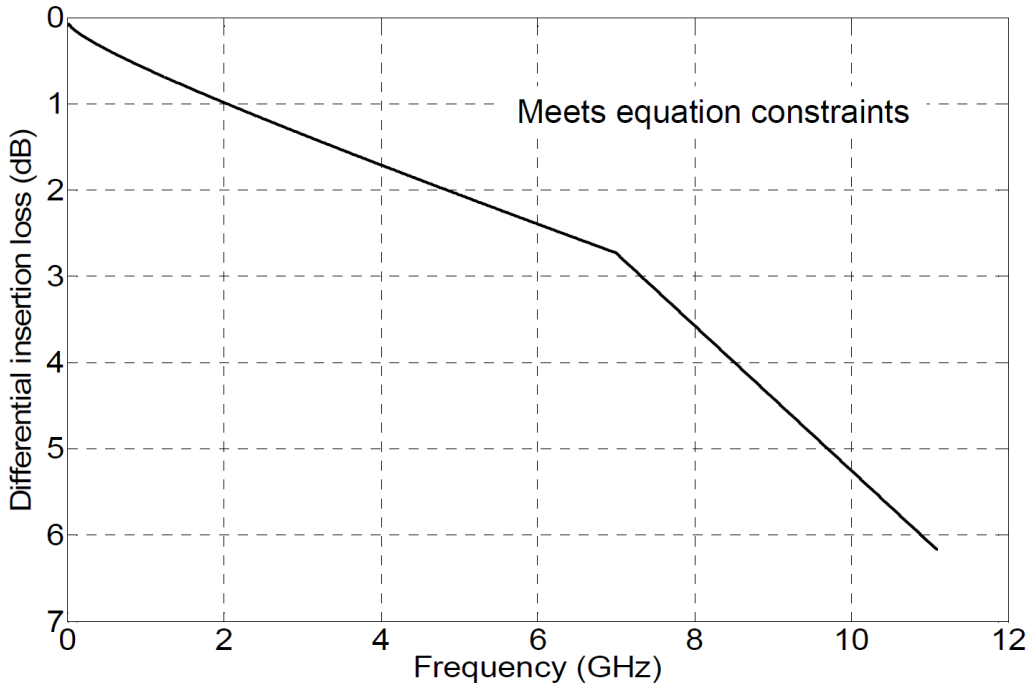


Figure I. 2. *Module insertion loss.*

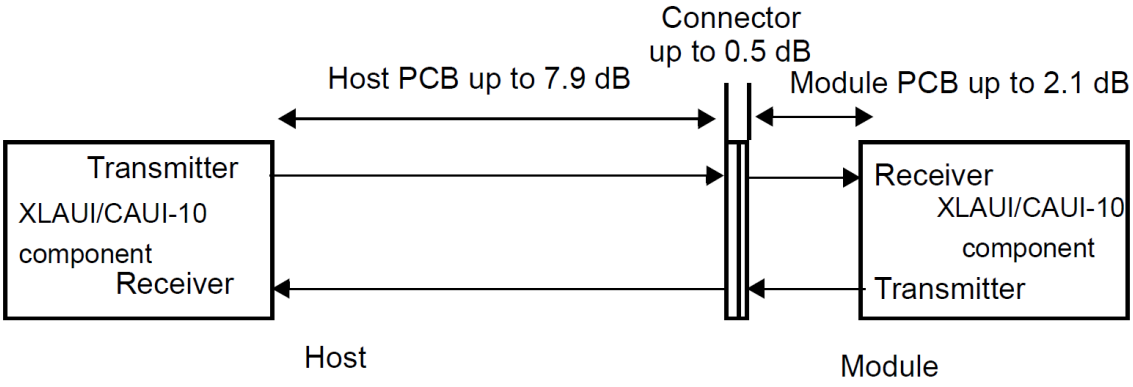


Figure I. 3. *Chip-to-module loss budget at 5.15625GHz.*

Section	Differential insertion loss max. (at 5.15625GHz)
Host CAUI-10 component to connector	7.9dB
Connector loss	0.5dB
Connector to module CAUI-10 component	2.1dB

Table I. 1. *Chip-to-module loss budget.*

I.1 Compliance point specifications for chip-to-module XLAUI/CAUI-10

The chip-to-module CAUI-10 interface specifies compliance points around the module connector as depicted in **Figure I. 5** and **Figure I. 7**. Chip-to-module devices shall meet the electrical characteristics defined in the Sections below. Compliance points are defined to ensure interoperability between hosts and modules. A Module Compliance Board (MCB) is used to isolate the characteristics of the module, and a Host Compliance Board (HCB) is used to isolate the characteristics of the host. **Figure I. 5** and **Figure I. 7** include the loss associated with the HCB and MCB at 5.15625GHz.

The reference differential insertion loss of the HCB PCB is given in **Equation I- 3** and illustrated in **Figure I. 4**. The effects of differences between the insertion loss of an actual HCB and the reference insertion loss are to be accounted in the measurements.

$$Insertion\_loss(f) = 0.017 + 0.5\sqrt{f} + 0.1836f \quad 0.01 \le f < 11.1$$

Equation I- 3

where

Insertion_loss(f) is the differential insertion loss at frequency *f*

f is the frequency in GHz

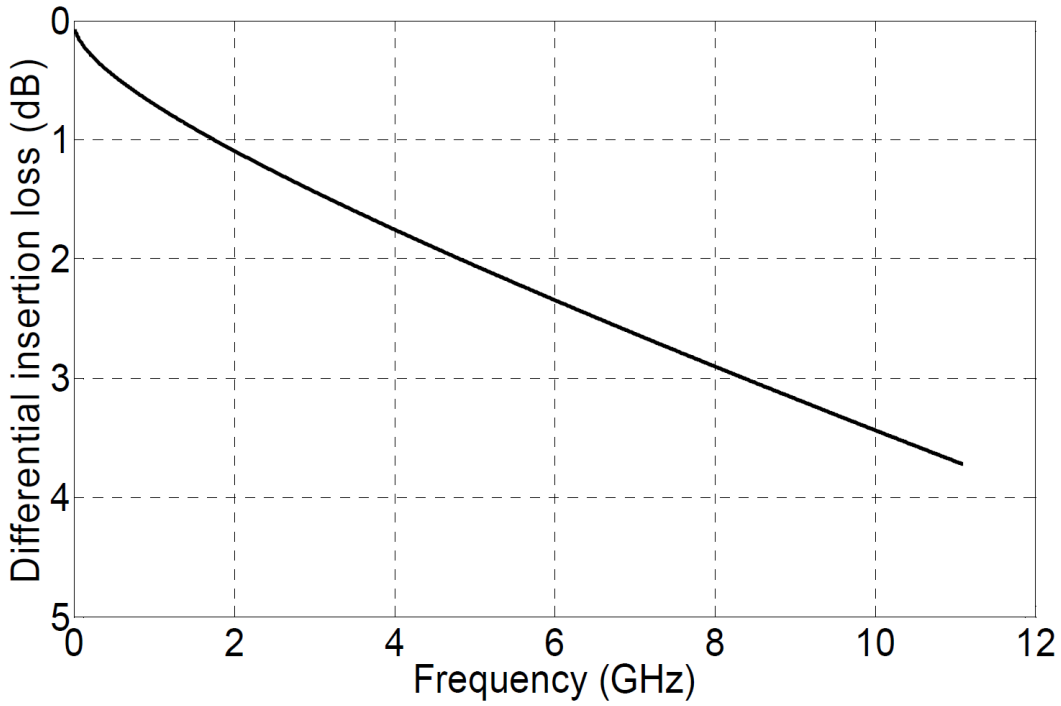


Figure I. 4. *HCB insertion loss.*

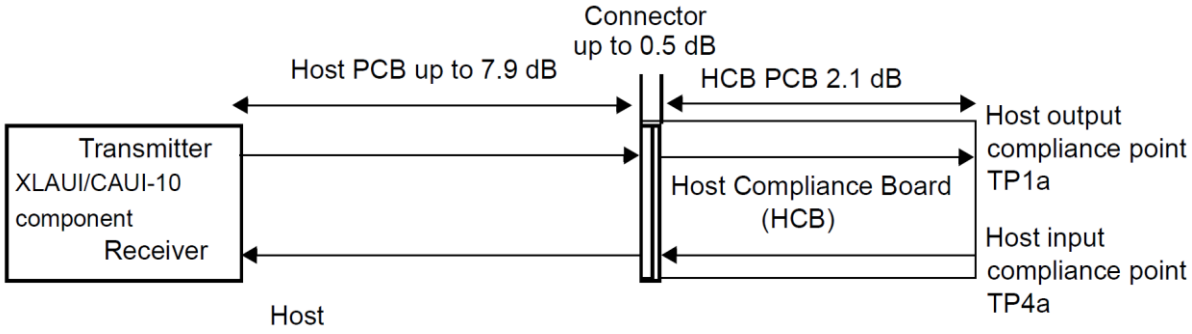


Figure I. 5. *Chip-to-module HCB insertion loss budget at 5.15625GHz.*

The reference differential insertion loss of the MCB PCB is given in **Equation I- 4** and illustrated in **Figure I. 6**. The effects of differences between the insertion loss of an actual MCB and the reference insertion loss are to be accounted in the measurements.

$$Insertion_{loss(f)} = -0.00086 + 0.2286\sqrt{f} + 0.08386f \qquad 0.01 \le f < 11.1$$

Equation I- 4

where

Insertion_loss(f) is the differential insertion loss at frequency *f*

f is the frequency in GHz

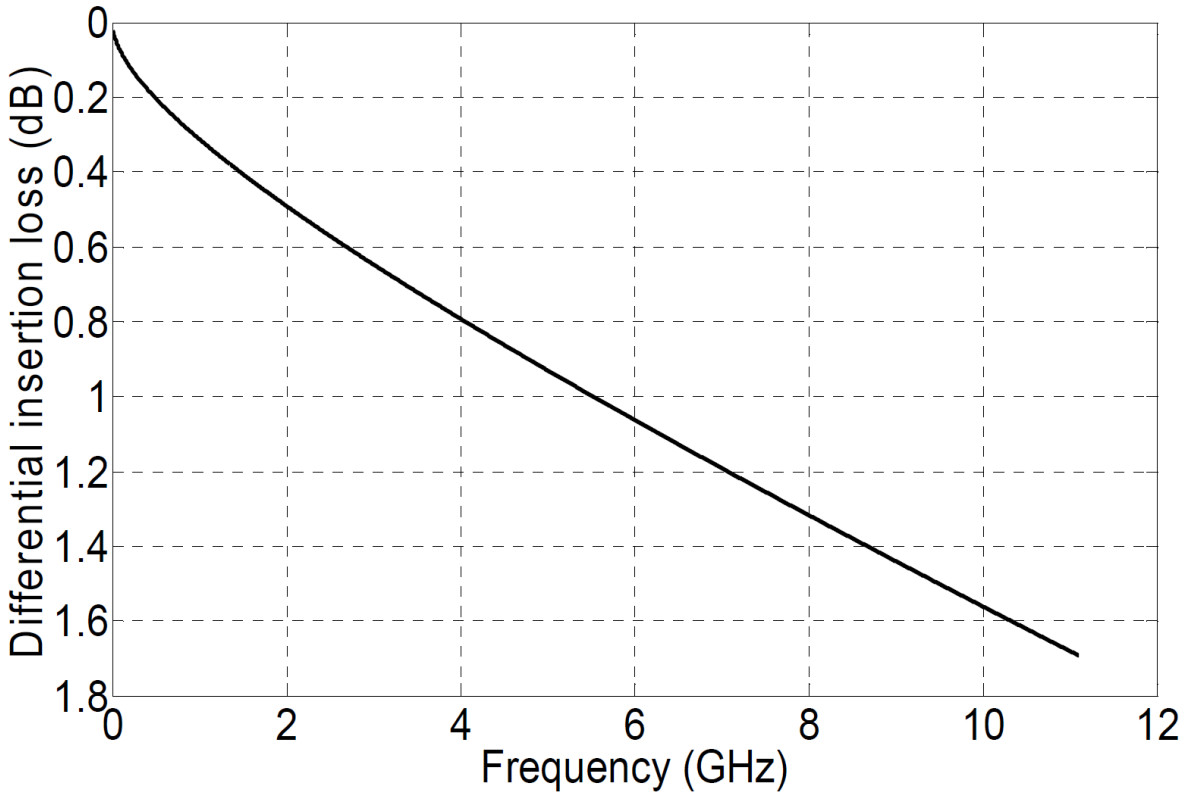


Figure I. 6. *MCB insertion loss.*

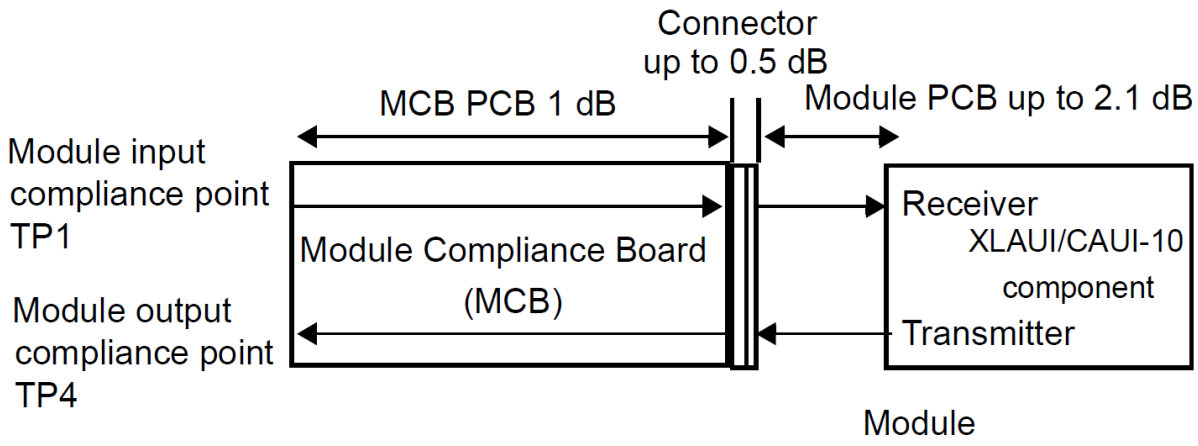


Figure I. 7. *Chip-to-module with MCB insertion loss budget at 5.15625GHz.*

I.1.1 Module specifications

A module that uses CAUI-10 to interface with a host shall meet the characteristics outlined in **Table I. 2** and **Table I. 3** when measured using the MCB and HCB (where the HCB is used to calibrate inputs to the module). **Table I. 2** also lists the equivalent test points for the CPPI (see **Figure 2-14**).

Modules may support additional de-emphasis states, but the specification of additional states is outside the scope of this standard. De-emphasis shall be off during eye mask and jitter testing. Module electrical output de-emphasis off state is the optimal setting for module electrical output jitter and eye mask evaluation. AC-coupling for both Tx and Rx paths shall be located in the module.

Reference	Compliance point		Value	Unit
Minimum module differential input return loss	MCB input	TP1	See **Equation I- 5**	dB
Module input tolerance signal	HCB output	TP1a	See **Annex H**	
Module output signal	MCB output	TP4	See **Table I. 3**	
Minimum module differential output return loss	MCB output	TP4	See **Equation I- 5**	dB

Table I. 2. *Specifications at module compliance points*

$$Return\_loss(f) \geq \begin{cases} 12 - 2\sqrt{f} & 0.01 \leq f < 2.19 \\ 5.56 - 8.7log_{10}\left(\frac{f}{5.5}\right) & 2.19 \leq f \leq 11.1 \end{cases} (dB)$$

Equation I- 5

where

Return_loss(f) is the differential input return loss at frequency *f*

f is the frequency in GHz

Minimum module differential input/output return loss is illustrated in **Figure I. 8**.

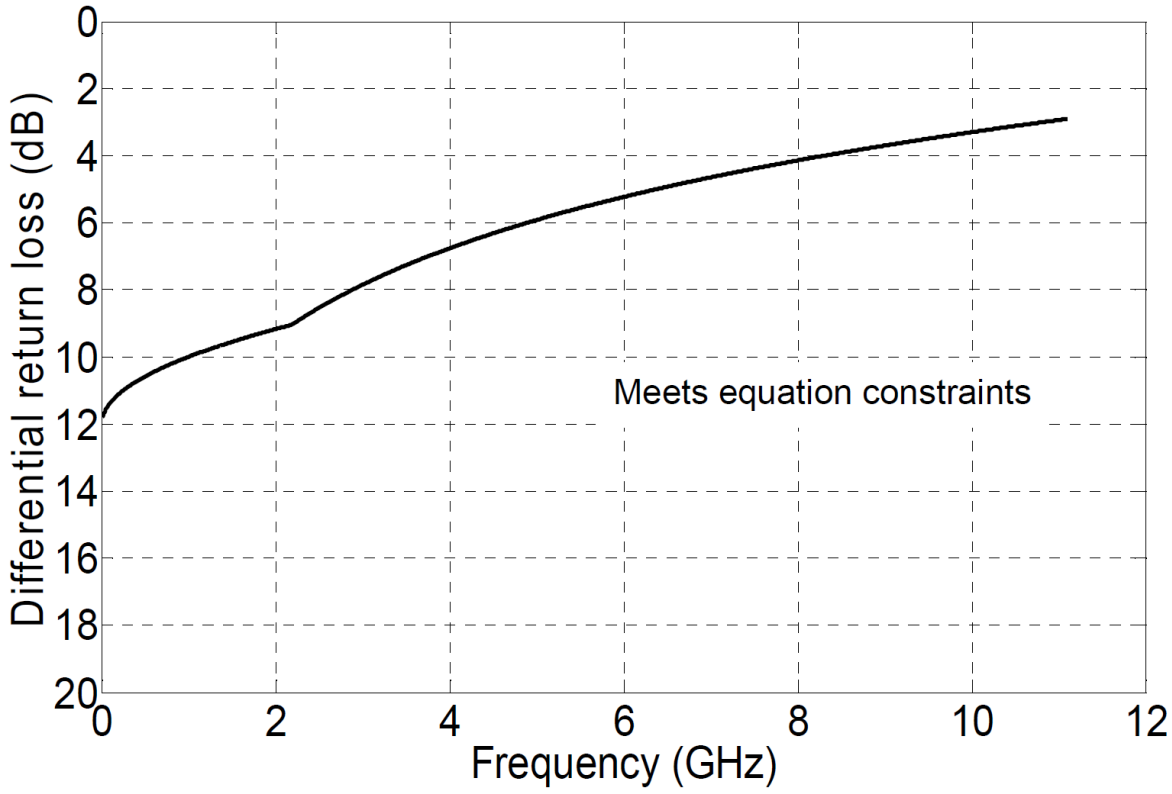


Figure I. 8. *Module input/output return loss.*

Parameter	Value	Unit
Maximum differential output voltage, peak-to-peak	760	mV
Minimum de-emphasis	3.5	dB
Maximum de-emphasis	6	dB
Minimum VMA	See **Equation I- 6**	mV
Maximum termination mismatch at 1 MHz	5	%

Maximum output AC common-mode voltage, RMS	15	mV
Minimum output rise and fall time (20% to 80%)	24	ps
Maximum Total Jitter	0.4	UI
Maximum Deterministic Jitter	0.25	UI
Module electrical output eye mask definition X1	0.2	UI
Module electrical output eye mask definition X2	0.5	UI
Module electrical output eye mask definition Y1	136	mV
Module electrical output eye mask definition Y2	380	mV

Table I. 3. *Module electrical output.*

$$Minimum\ VMA(mV) = (-110 - 2.13x + 0.32x^2) \times (10^{-y/20})$$

Equation I- 6

where

x is the rise or fall time, or 38, (whichever is larger) in ps

y is de-emphasis value in dB

I.1.2. Host specifications

A host that uses CAUI-10 to interface with a module shall meet the characteristics outlined in **Table I. 4** and **Table I. 5** when measured using the HCB and MCB (where MCB is used to calibrate inputs to the host). **Table I. 4** also lists the equivalent test points for the CPPI.

Reference	Compliance point		Value	Unit
Host output signal specifications	HCB output	TP1a	See **Table I. 5**	
Minimum host differential output return loss	HCB output	TP1a	See **Equation I- 7**	dB
Minimum host differential input return loss	HCB input	TP4a	See **Equation I- 8**	dB
Host input tolerance signal	MCB output	TP4	See Section **I.1.3**	

Table I. 4. *Specifications at host compliance points*

$$Return\_loss(f) \geq \begin{cases} 12 - 2\sqrt{f} & 0.01 \leq f < 2.19 \\ 5.56 - 8.7log_{10}\left(\frac{f}{5.5}\right) & 2.19 \leq f \leq 11.1 \end{cases} (dB)$$

Equation I- 9

where

Return_loss(f) is the differential input return loss at frequency *f*

f is the frequency in GHz

Minimum host differential input/output return loss is illustrated in **Figure I. 9**.

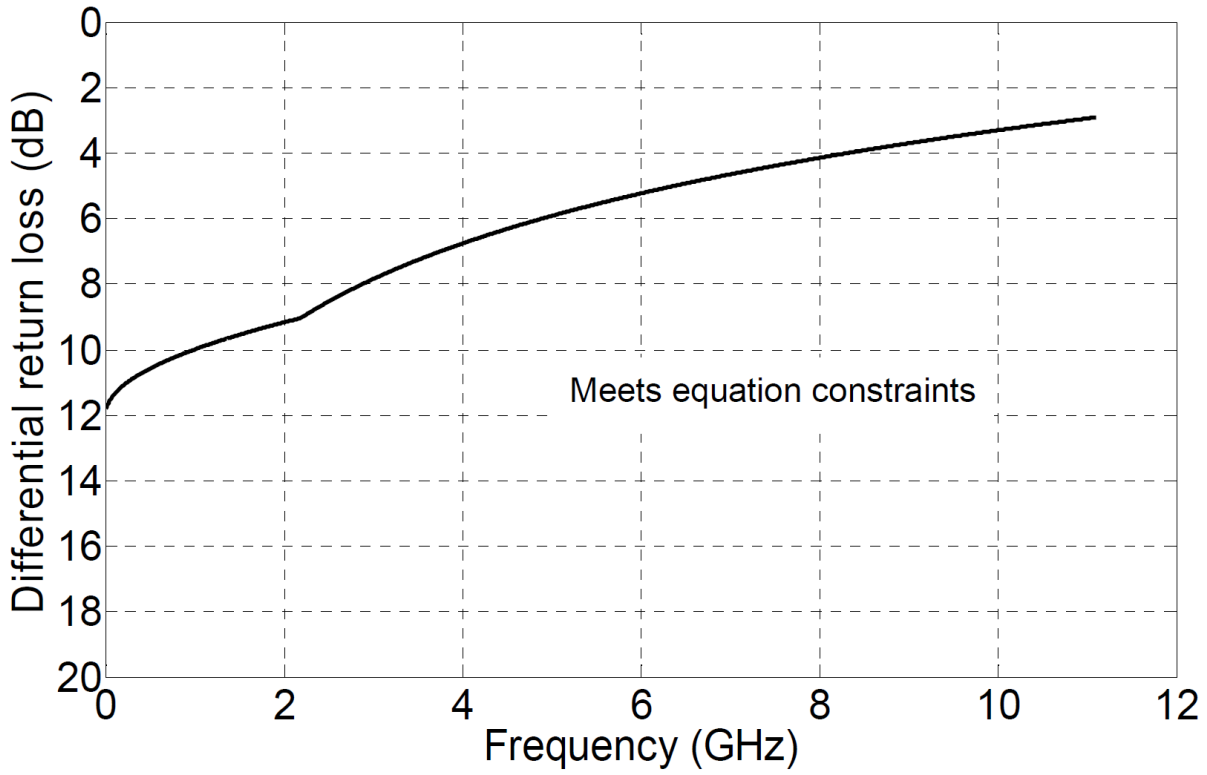


Figure I. 9. *Host input/output return loss.*

Parameter	Value	Units
Maximum output AC common-mode voltage, RMS	20	mV
Minimum output rise and fall time (20% to 80%)	24	ps
Maximum Total Jitter	0.62	UI
Maximum Deterministic Jitter	0.42	UI
Host electrical output eye mask definition X1	0.31	UI
Host electrical output eye mask definition X2	0.5	UI
Host electrical output eye mask definition Y1	42.5	mV
Host electrical output eye mask definition Y2	42.5	mV

Table I. 5. *Host electrical output*

I.1.3. Host input signal tolerance

Host CAUI-10 jitter tolerance evaluation shall be defined by a stressed input signal that comprises 0.25 UI Deterministic Jitter, and 0.15 UI random jitter for BER 10^{-12}. Deterministic Jitter is added to a clean test pattern as sinusoidal jitter defined in **Annex H**. The limited low-pass filter stress is added until the 0.22 UI Deterministic Jitter is achieved. Frequency-dependent attenuation is then added using PCB trace or frequency-dependent attenuation that emulates PCB loss. Frequency-dependent attenuation is added until 0.25 UI Deterministic Jitter is achieved. Random jitter is added to the test signal using an interference generator, which is a broadband noise source capable of producing white Gaussian noise with adjustable amplitude. The power spectral density shall be flat to ±3dB from 50MHz to 6GHz with a crest factor of no less than 5. **Figure I. 10** depicts a CAUI-10 jitter tolerance test setup. The amplitude and output jitter of the filter stress plus limiter and random jitter injection shall meet the receiver eye mask illustrated in **Figure I. 11** with the values for X1, X2, Y1, Y2 given in **Table I. 3**. All CAUI-10 lanes shall be active during jitter tolerance testing.

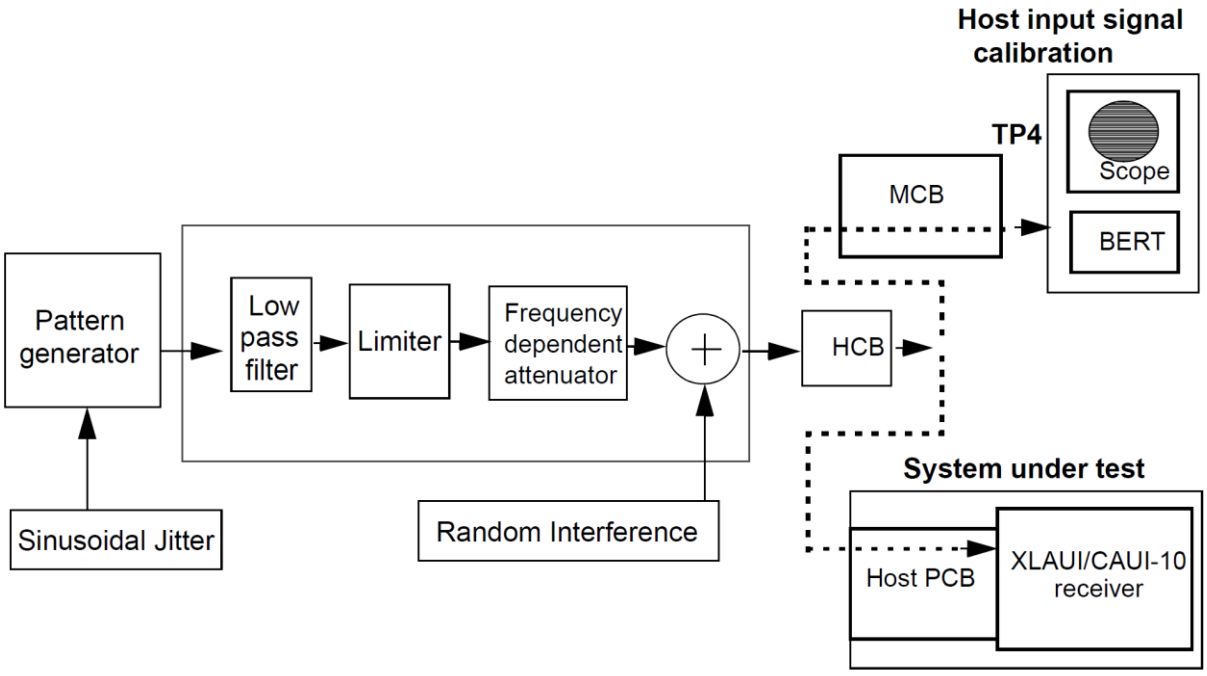


Figure I. 10. *Stressed-eye and jitter tolerance test setup.*

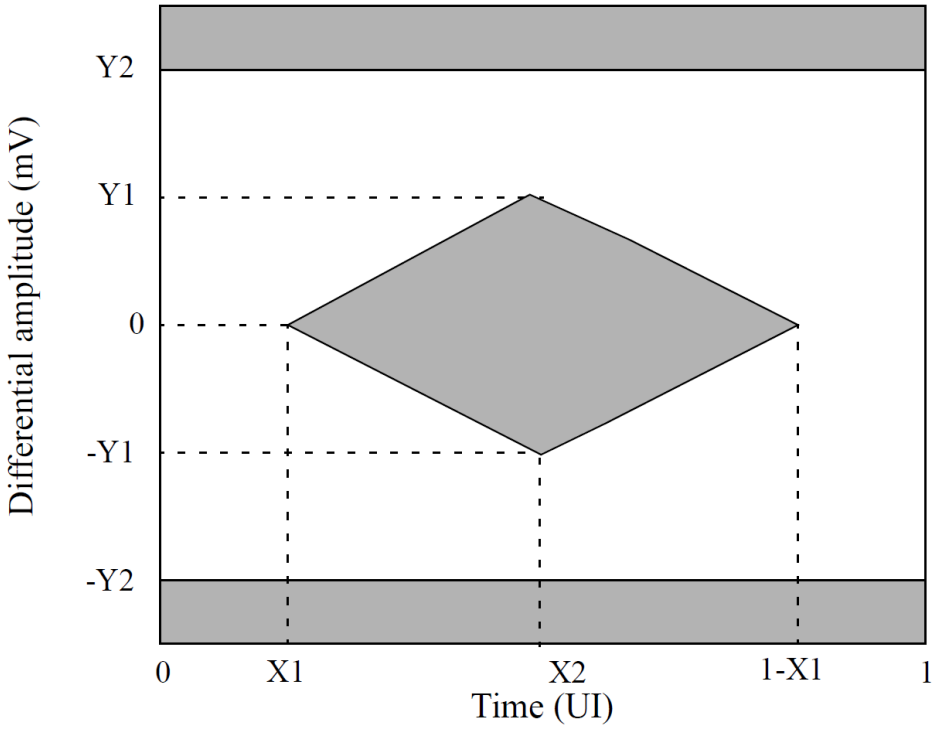


Figure I. 11. *Receiver template.*

Annex J. PMA sublayer partitioning examples

The following subclauses provide various partitioning examples. Partitioning guidelines and MMD numbering conventions are described in Section **2.5**.

J.1. Partitioning examples with FEC

The example of BASE-R FEC implemented in a separate device from either the PCS or the PMD is illustrated in **Figure J. 2**.

J.1.1. FEC implemented with PCS

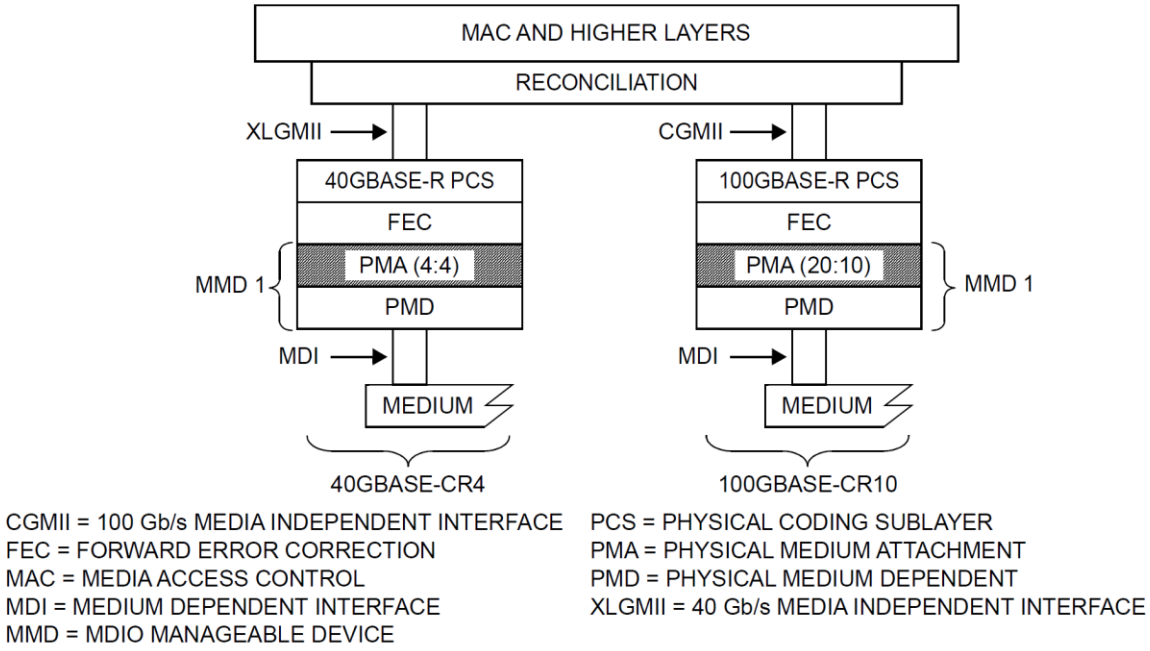


Figure J. 1. *Example FEC implemented with PCS*

J.1.2. FEC implemented with PMD

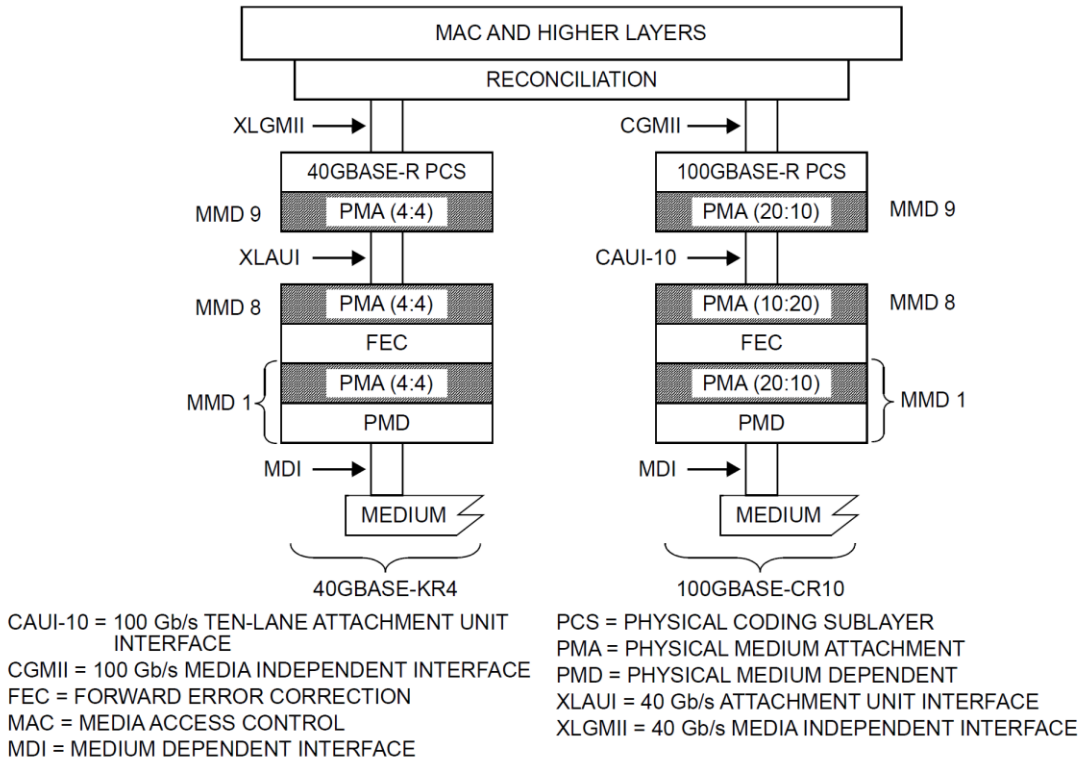


Figure J. 2. *Example FEC implemented with PMD*

J.2. Partitioning examples with RS-FEC

J.2.1. Single PMA sublayer with RS-FEC

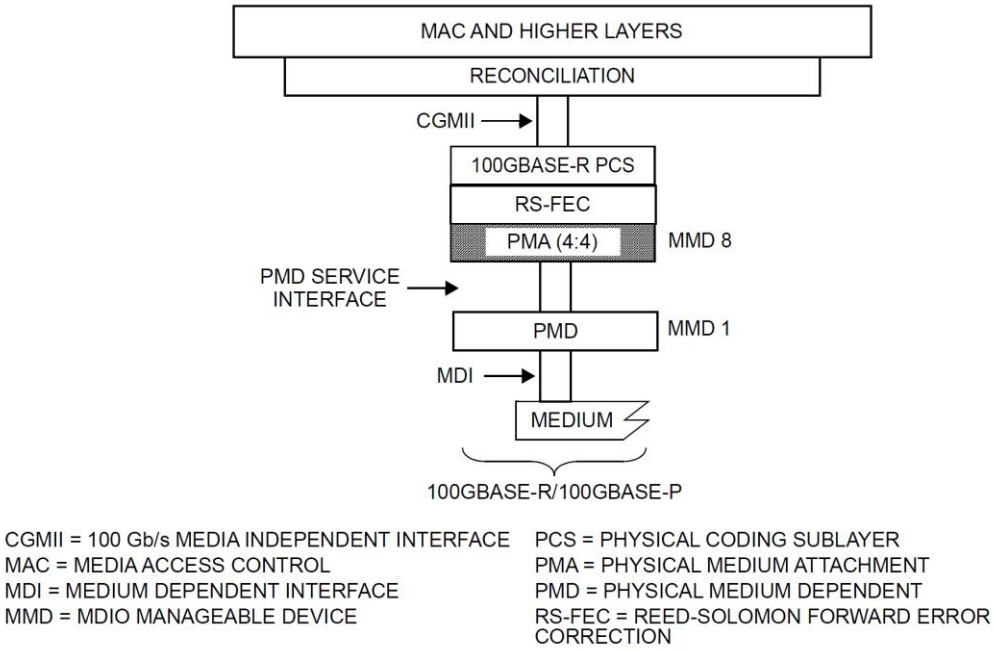


Figure J. 3. *Example single PMA sublayer with RS-FEC.*

J.2.2. Single CAUI-10 with RS-FEC

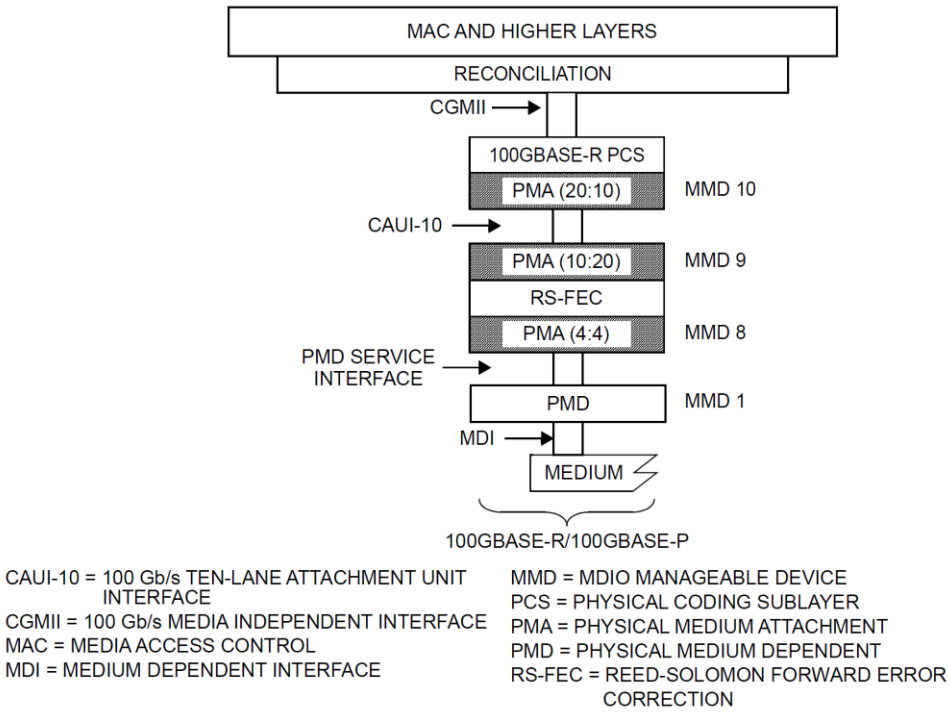


Figure J. 4. *Example single CAUI-10 with RS-FEC.*

J.3. Partitioning examples without FEC

J.3.1. Single PMA sublayer without FEC

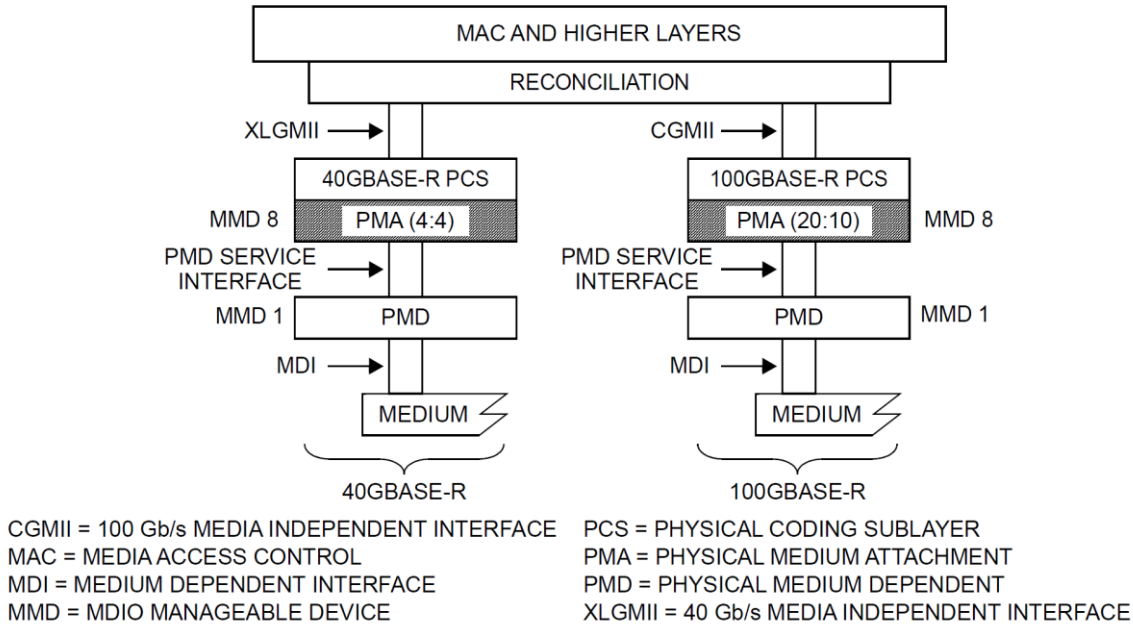


Figure J. 5. *Example single PMA sublayer without FEC.*

J.3.2. Single XLAUI/CAUI-4 without FEC

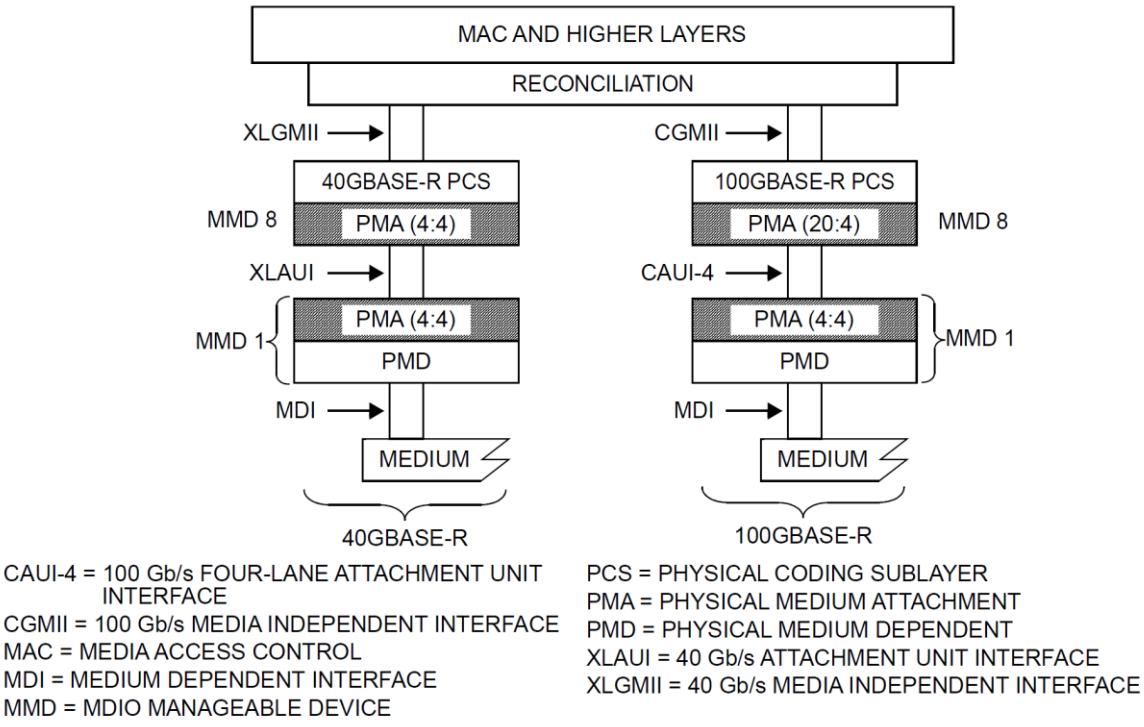


Figure J. 6. *Example single XLAUI/CAUI-4 without FEC.*

J.3.3. Separate SERDES for optical module interface

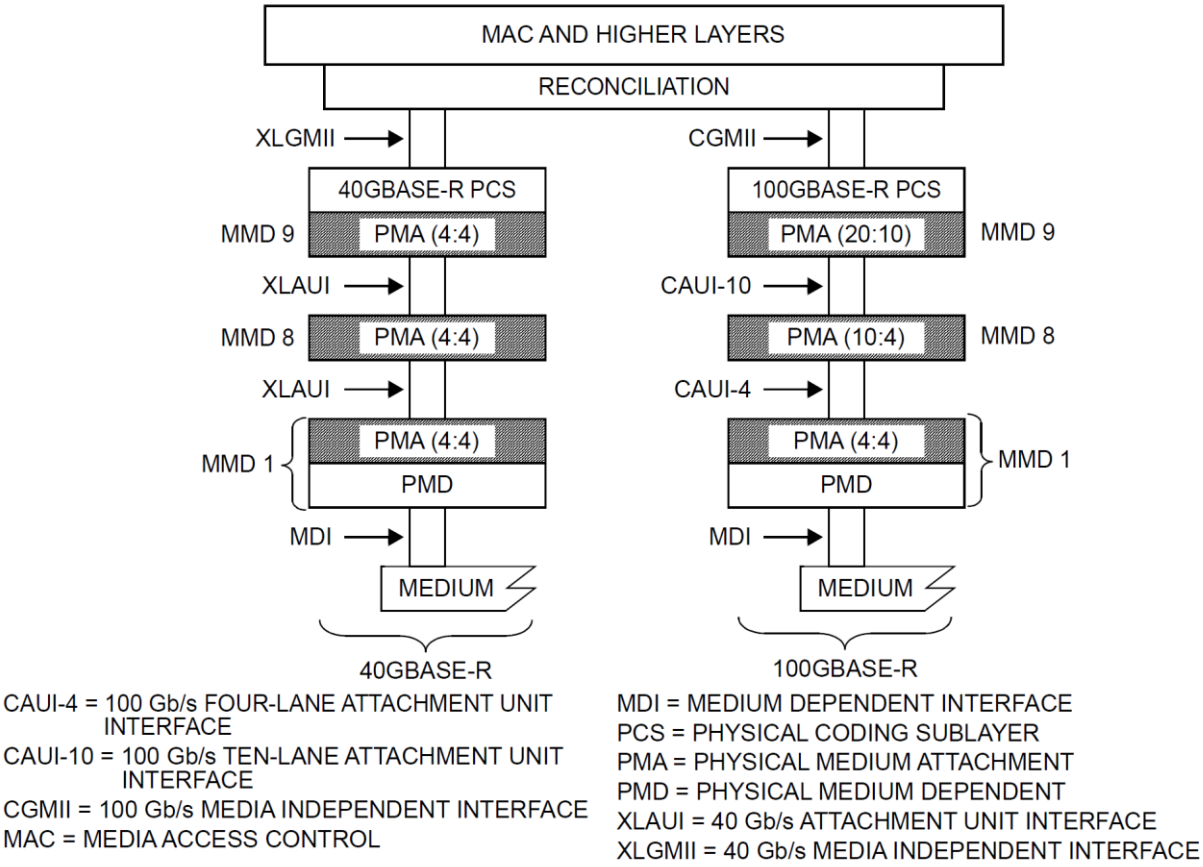


Figure J. 7. *Separate SERDES for optical module interface.*

Annex K. Chip-to-chip 100Gb/s four-lane Attachment Unit Interface (CAUI-4)

This annex defines the functional and electrical characteristics for the optional chip-to-chip 100 Gb/s four-lane Attachment Unit Interface (CAUI-4). **Figure K. 1** shows an example relationship of the CAUI-4 C2C interface to the ISO/IEC OSI reference model. The C2C interface provides the electrical characteristics and the associated compliance points, which can optionally be used when designing systems with electrical interconnect of approximately 25cm in length.

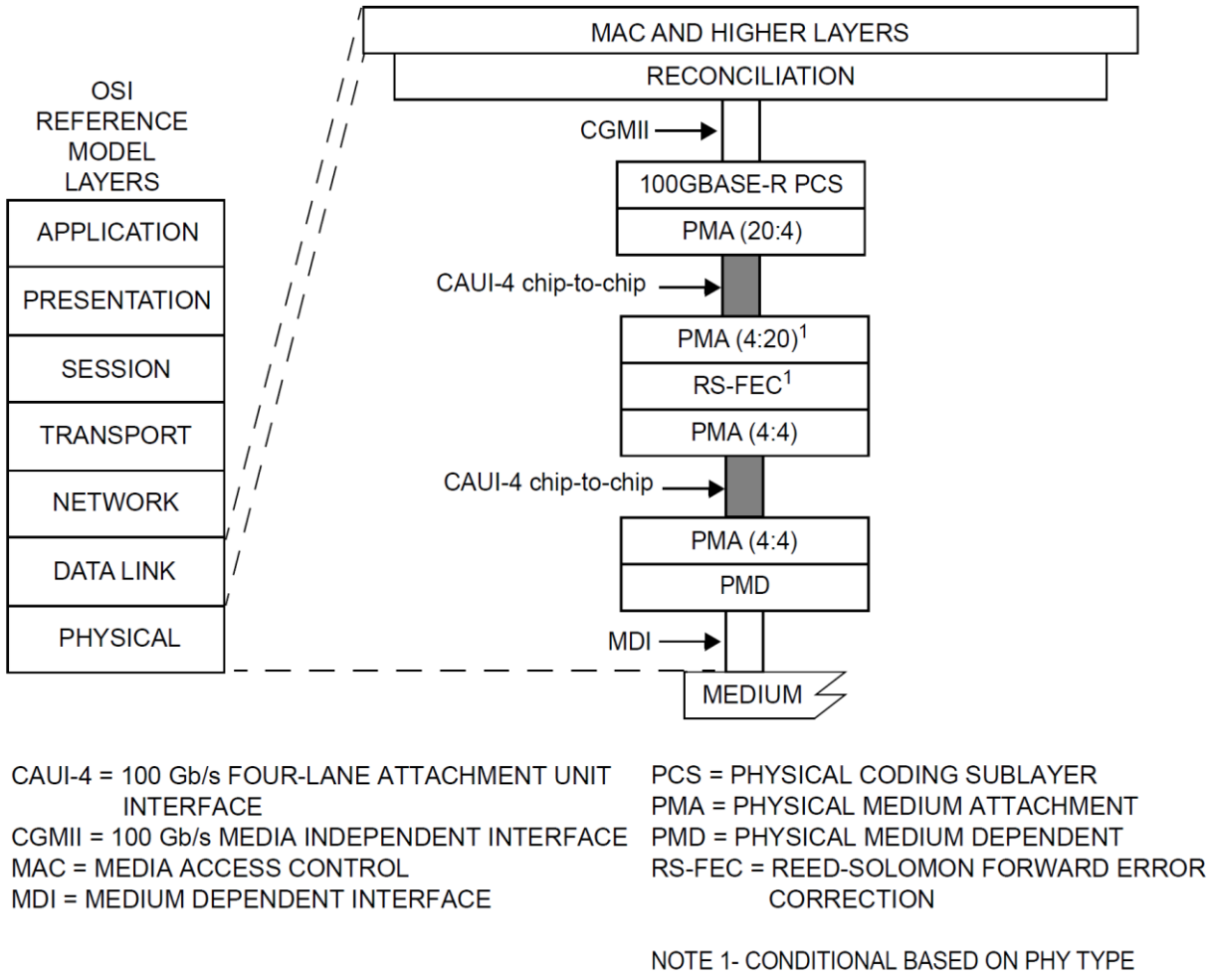


Figure K. 1. *Example CAUI-4 chip-to-chip relationship to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model.*

The CAUI-4 bidirectional link comprises of a CAUI-4 transmitter, a CAUI-4 channel, and a CAUI-4 receiver. **Figure K. 2** depicts a typical CAUI-4 application, and **Equation K- 1**

(illustrated in **Figure K. 3**) summarizes the differential insertion loss budget associated with the C2C application. The CAUI-4 C2C interface represents the independent data paths in each direction. Each data path contains four differential lanes, which are AC-coupled. The nominal signaling rate of each lane is 25.78125GBd. The CAUI-4 transmitter on each end of the link is adjusted to an appropriate setting based on the channel. If implemented, the feedback mechanism of the transmitter equalization, which is described below, may be used to identify an appropriate setting. The adaptive or adjustable receiver reacts as the remainder of the equalization.

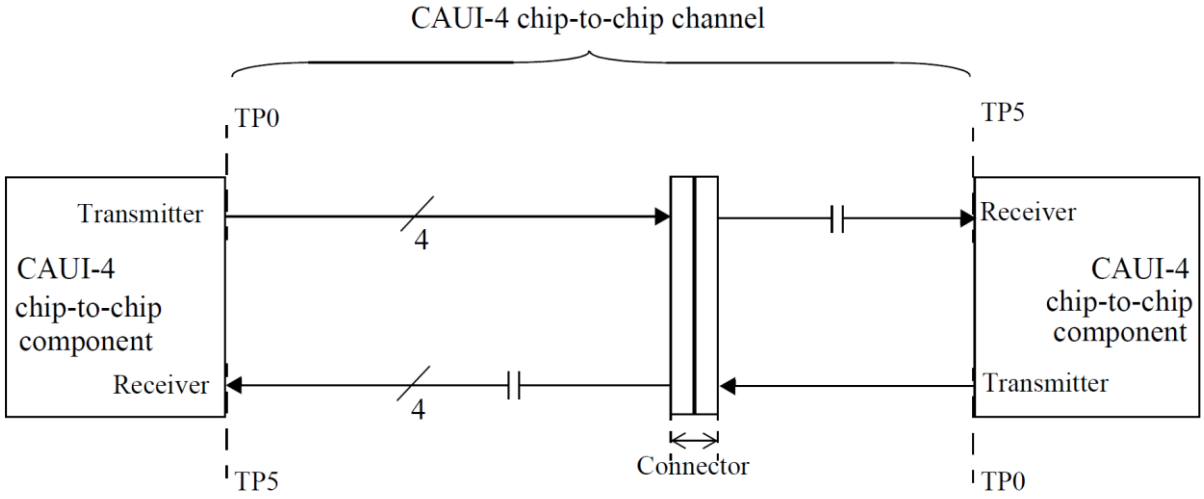


Figure K. 2. *Typical CAUI-4 chip-to-chip application.*

The normative channel compliance is through CAUI-4 COM as described below. Actual channel loss could be higher or lower than that given by **Equation K- 1** due to the channel ILD, return loss, and crosstalk.

$$Insertion\_loss(f) \leq \begin{Bmatrix} 1.083 + 2.543\sqrt{f} + 0.761f & 0.01 \leq f < 12{,}89 \\ -17.851 + 2.936f & 12{,}89 \leq f \leq 25.78 \end{Bmatrix} \ (dB)$$

Equation K- 1

where

Insertion_loss(f) is the informative CAUI-4 chip-to-chip insertion loss

f is the frequency in GHz

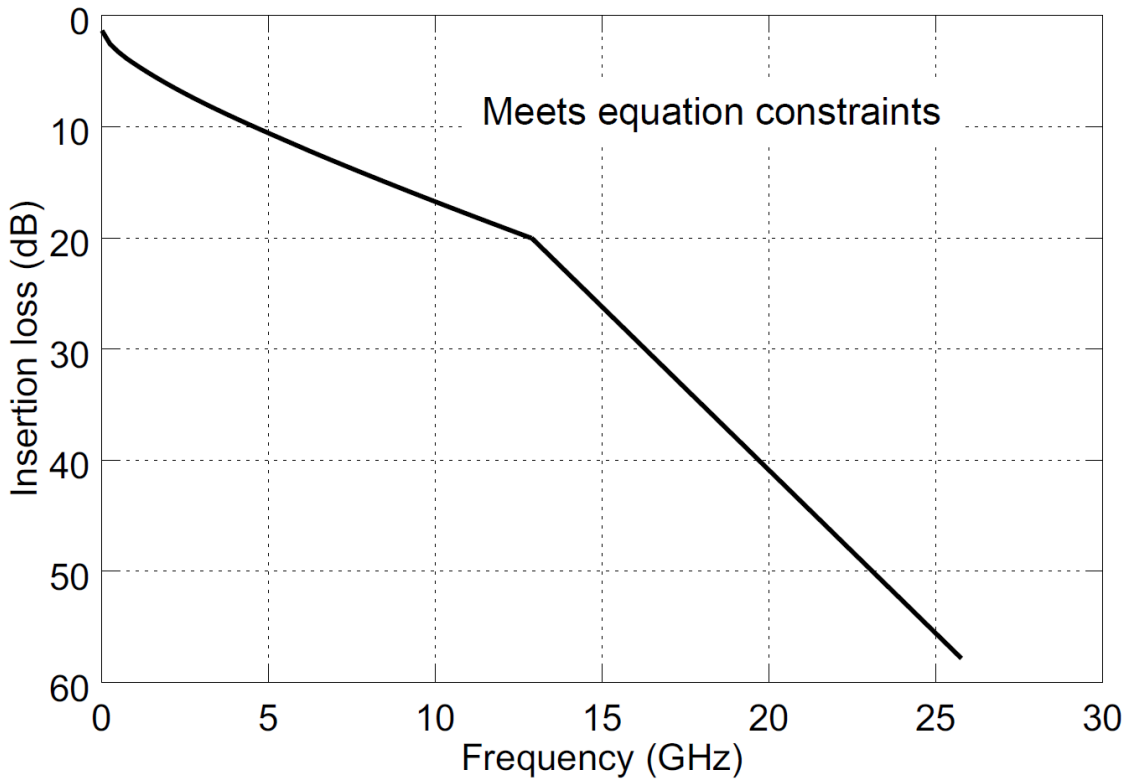


Figure K. 3. *CAUI-4 chip-to-chip channel insertion loss*

K.1. CAUI-4 chip-to-chip compliance point definition

The electrical characteristics for the CAUI-4 chip-to-chip interface are defined at compliance points for the transmitter (TP0a) and receiver (TP5a), respectively.

K.2. CAUI-4 chip-to-chip electrical characteristics

K.2.1. CAUI-4 transmitter characteristics

A CAUI-4 chip-to-chip transmitter shall meet the specifications defined in **Table K. 1** if measured at TP0a. While the CAUI-4 chip-to-chip transmitter requirements are similar to 100GBASE-KR4, they differ in that they do not assume transmitter training or a back-channel communications path. Also, the transmit output waveform is not manipulated via the PMD control function but may optionally be manipulated via the feedback mechanism described below.

A test system with a fourth-order Bessel-Thomson low-pass response with 33GHz 3dB bandwidth is to be used for all transmitter signal measurements, unless otherwise specified.

Parameter	Value	Units
Signaling rate per lane (range)	25.78125 ± 100 ppm	GBd
Differential peak-to-peak output voltage (max)		
- Transmitter disabled	30	mV
- Transmitter enabled	1200	mV
Common-mode voltage (max)	1.9	V
Common-mode voltage (min)	0	V
AC common-mode output voltage (max, RMS)	12	mV
Differential output return loss (min)	$RL_d(f) \geq \begin{cases} 12.05 - f & 0.05 \leq f \leq 6 \\ 6.5 - 0.075f & 6 < f \leq 19 \end{cases}$	dB
Common-mode output return loss (min)	$RL_{cm}(f) \geq \begin{cases} 9.05 - f & 0.05 \leq f \leq 6 \\ 3.5 - 0.075f & 6 < f \leq 19 \end{cases}$	dB
Output waveform [a]		
- Steady state voltage v_f (max)	0.6	V
- Steady state voltage v_f (min)	0.4	V
- Linear fit pulse peak (min)	0.71 × v_f	V
- Pre-cursor equalization	**Table K. 2**	-
- Post-cursor equalization	**Table K. 3**	-
Signal-to-noise-and-distortion ratio, SNDR (min)	27	dB
Output Jitter (max)		
- Even-odd jitter	0.035	UI
- Effective bounded uncorrelated jitter, peak-to-peak [b]	0.1	UI
- Effective total uncorrelated jitter, peak-to-peak [b, c]	0.26	UI

[a] The state of the transmit equalizer is controlled by management interface.

[b] Effective bounded uncorrelated jitter and effective total uncorrelated jitter shall be from 10^{-6} to 10^{-4}.

[c] Effective total uncorrelated jitter, peak-to-peak is specified to a 10^{-15} probability.

Table K. 1. *CAUI-4 transmitter characteristics at TP0a.*

K.2.1.1. Transmitter equalization settings

The CAUI-4 C2C transmitter includes programmable equalization to compensate for the frequency-dependent loss of the channel and to facilitate data recovery at the receiver. The functional model for the transmit equalizer is the three-tap transversal filter shown in **Figure K. 4.** Transmit equalizer functional model. The transmitter output equalization is characterized using the linear fit method where the state of the CAUI-4 transmit output is manipulated via management.

The variable *Local_eq_cm1* controls the weight of the pre-cursor tap *c(–1)*, by changing the ratio *c(–1)/(|c(–1)|+|c(0)|+|c(1)|)*. The valid values of *Local_eq_cm1* variable and the corresponding ratios are specified in **Figure K. 2.** Typical CAUI-4 chip-to-chip application. The variable *Local_eq_c1* controls the weight of the post-cursor tap *c(1)*, by changing the ratio *c(1)/(|c(–1)|+|c(0)|+|c(1)|)*. The valid values of *Local_eq_c1* and the corresponding ratios are specified in **Table K. 3**. The *Local_eq_cm1* and *Local_eq_c1* variables are independent of each other and independent on each lane. Each successive step in the value of the *Local_eq_cm1* and *Local_eq_c1* variables shall result in a monotonic change in transmitter equalization.

***Local_eq_cm1* value**	$\boldsymbol{c(-1)ratio\left(\frac{c(-1)}{\lvert c(-1)\rvert+\lvert c(0)\rvert+\lvert c(1)\rvert}\right)}$
0	0 ±0.04
1	–0.05 ±0.04
2	–0.1 ±0.04
3	–0.15 ±0.04

Table K. 2. *Pre-cursor equalization.*

***Local_eq_cm1* value**	$\boldsymbol{c(-1)ratio\left(\frac{c(-1)}{\lvert c(-1)\rvert+\lvert c(0)\rvert+\lvert c(1)\rvert}\right)}$
0	0 ±0.04
1	–0.05 ±0.04
2	–0.1 ±0.04
3	–0.15 ±0.04
4	–0.2 ±0.04
5	–0.25 ±0.04

Table K. 3. *Post-cursor equalization.*

If an MDIO is implemented, the *Local_eq_cm1* and *Local_eq_c1* variables for each lane (0 through 3) and direction (transmit and receive) are accessible through the registers 1.180 to 1.187.

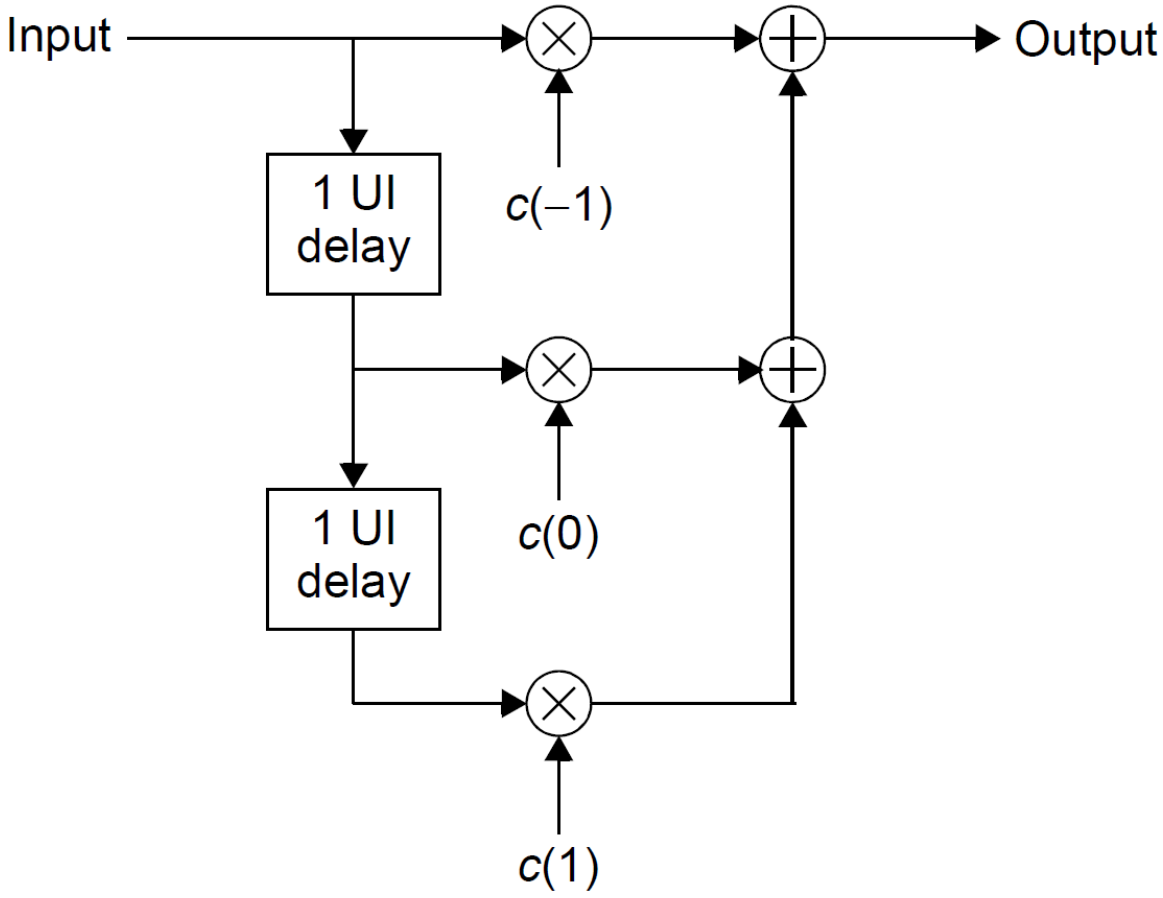


Figure K. 4. *Transmit equalizer functional model.*

K.2.2. Optional Energy Efficient Ethernet (EEE) operation

If the optional EEE capability with the deep sleep mode option is supported (see **Annex F**), then the inter-sublayer service interface includes four additional primitives as described in 83.3 and may also support CAUI-4 shutdown.

If the EEE capability includes CAUI-4 shutdown (see **Annex F**), then when *aui_tx_mode* is set to *ALERT*, the transmit direction sublayer sends a repeating 16-bit pattern, hexadecimal 0xFF00, which is transmitted across the CAUI-4. This sequence is transmitted regardless of the value of the *tx_bit*, which is presented by the *PMA:IS_UNITDATA_i.request* primitive or thevalue of the *rx_bit*, which is presented by

the *PMA:IS_UNITDATA_i.indication* primitive. When the *aui_tx_mode* is *QUIET*, the transmit direction of the CAUI-4 transmitter is disabled, as specified below. When the received *aui_tx_mode* is set to *ALERT*, the receive direction sublayer sends a repeating 16-bit pattern (hexadecimal 0xFF00), which is transmitted across the CAUI-4. This sequence is transmitted regardless of the value of the *tx_bit*, which is presented by the *PMA:IS_UNITDATA_i.request* primitive or the *rx_bit, which is* presented by the *PMA:IS_UNITDATA_i.indication* primitive. When the received *aui_tx_mode* is set to *QUIET*, the receive direction CAUI-4 transmitter is disabled as specified below.

For EEE capability with the CAUI-4 shutdown, the differential peak-to-peak output voltage of the CAUI-4 transmitter lane's shall be less than 30mV within 500ns of the *aui_tx_mode* changing to *QUIET* in the relevant direction. Furthermore, the CAUI-4 transmitter lane's differential peak-to-peak output voltage shall be greater than 720mV within 500ns of *aui_tx_mode* ceasing to be *QUIET* in the relevant direction.

Global transmit disable is optional for EEE capability. The transmit disable function shall turn off all transmitter lanes for a physically instantiated AUI in either the ingress or the egress direction. In the egress direction, the PMA turns off all the transmitter lanes for the CAUI-4 egress direction, if *PEASE* is asserted and the *aui_tx_mode* is set to *QUIET*. In the ingress direction, the PMA turns off all the transmitter lanes for the CAUI-4 ingress direction, if PEASE is asserted and the received *aui_tx_mode* is set to *QUIET*. In both directions, the transmit disable function shall turn on all transmitter lanes after the appropriate direction *aui_tx_mode* changes to any state other than *QUIET* within a time and voltage level specified in this annex.

K.2.3. CAUI-4 receiver characteristics

A CAUI-4 chip-to-chip receiver shall meet the specifications defined in **Table K. 4** if measured at TP5a.

Parameter	Value	Units
Differential input return loss (min)	$RL_d(f) \geq \begin{cases} 12.05 - f & 0.05 \leq f \leq 6 \\ 6.5 - 0.075f & 6 < f \leq 19 \end{cases}$	dB
Differential to common-mode input return loss	$RL_{cd}(f) \geq \begin{cases} 25 - 1.44f & 0.05 \leq f \leq 6.95 \\ 15 & 6.95 < f \leq 19 \end{cases}$	dB
Interference tolerance	**Table K. 5**	-

Table K. 4. *CAUI-4 receiver characteristics at TP5a.*

K.2.3.1. Receiver interference tolerance

The receiver shall satisfy the requirements for interference tolerance defined in **Table K. 5**. The interference tolerance test uses the following exceptions:

a. The parameters in **Table K. 5**.
b. The transmitter taps are set via management (see below) to the settings that provide the lowest BER.
c. Sinusoidal jitter is added to the test transmitter by modulating the clock source.

Parameter	Test 1 values			Test 2 values			Units
	Min	Max	Target	Min	Max	Target	
Bit error ratio [a, b]	-	10^{-15}		-	10^{-15}		-
Applied pk-pk sinusoidal jitter			**Table 2-32**			**Table 2-32**	
Insertion loss at 12.89 GHz [c]	19.5	20.5		9.5	10.5		dB
Coefficients of fitted insertion loss [d] a_0 a_1 a_2 a_4	 -1 0 0 0	 2 2.937 1.599 0.03		 -1 0 0 0	 1 0.817 0.801 0.01		 dB $dB/GHz^{1/2}$ dB/GHz dB/GHz^2
RSS_DFE2 [e]	0.05	-		0.025	-		-
COM including effects of broadband noise	-	-	2	-	-	2	dB

[a] BER replaces the RS symbol error ratio measurement.

[b] Maximum BER assumes errors are not correlated to ensure sufficiently high Mean Time To False Packet Acceptance (MTTFPA) assuming 64B/66B coding. Actual implementation of the receiver is beyond the scope of this standard.

[c] Measured between TP_t and TP5.

[d] Coefficients are calculated from the insertion loss measured between TP_t and TP5 with f_{min} = 0.05GHz, and $_{fmax}$ = 25.78125GHz, and maximum Δ_f = 0.01GHz.

[e] RSS_DFE2 is equivalent to RSS_DFE4 except that n_1=2 and n_2=5.

Table K. 5. *Receiver interference tolerance parameters.*

K.2.4. Global energy detect function for optional EEE operation

The global energy detect function is mandatory for EEE capability with the deep sleep mode option and CAUI-4 shutdown. The global energy detect function indicates whether or not signaling energy is being received on the physical instantiation of the inter sublayer interface (in each direction as appropriate). The energy detection function can be considered a subset of the signal indication logic. If no energy is being received on the CAUI-4 for the ingress direction, *SIGNAL_DETECT* is set to *FAIL* following a transition from *aui_rx_mode = DATA* to *aui_rx_mode = QUIET*. When *aui_rx_mode = QUIET*, *SIGNAL_DETECT* shall be set to *OK* within 500ns following the application of a signal at the receiver input that corresponds to an *ALERT* signal driven from the CAUI-4 link partner. While *aui_rx_mode = QUIET*, *SIGNAL_DETECT* changes from *FAIL* to *OK* only after the valid *ALERT* signal is received.

K.3. CAUI-4 chip-to-chip channel characteristics

The COM, computed using the parameters in **Table K. 6**, shall be greater than or equal to 2dB. This minimum value allocates margin for practical limitations on the receiver implementation, and the allowed transmitter equalization coefficients.

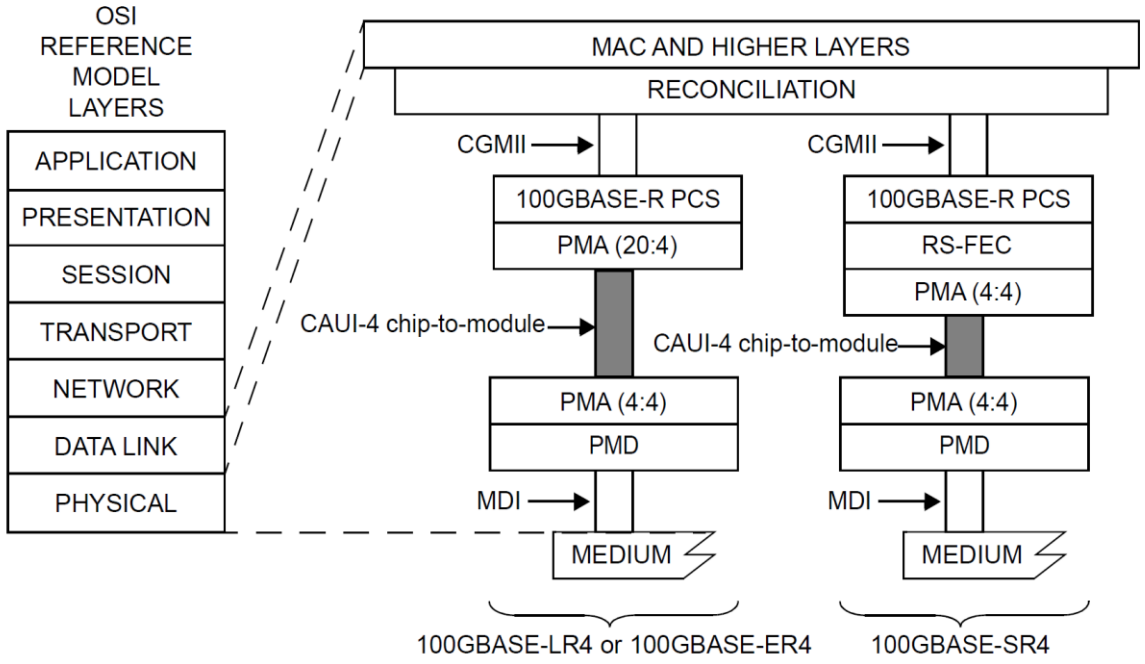
Parameter	Symbol	Value	Units
Signaling rate	f_b	25.78125	GBd
Maximum start frequency	f_{min}	0.05	GHz
Maximum frequency step	Δf	0.01	GHz
Device package model			
- Single-ended device capacitance	C_d	2.5×10^{-4}	nF
- Transmission line length, Test 1	z_p	12	mm
- Transmission line length, Test 2	z_p	30	mm
- Single-ended board capacitance	C_p	1.8×10^{-4}	nF
Single-ended reference resistance	R_0	50	ohms
Single-ended termination resistance	R_d	55	ohms
Receiver 3 dB bandwidth	f_r	$0.75 \times f_b$	GHz
Transmitter equalizer, minimum cursor coefficient	*c(0)*	0.6	-
Transmitter equalizer, pre-cursor coefficient	*c(–1)*		-

- Minimum value - Maximum value - Step size		-0.15 0 0.05	
Transmitter equalizer, post-cursor coefficient - Minimum value - Maximum value - Step size	*c(1)*	 -0.25 0 0.05	-
Continuous time filter, DC gain - Minimum value - Maximum value - Step size	g_{DC}	 -12 0 1	dB
Continuous time filter, zero frequency for $g_{DC} = 0$	f_z	$f_b/4$	GHz
Continuous time filter, pole frequencies	f_{p1} f_{p2}	$f_b/4$ f_b	GHz
Transmitter differential peak output voltage - Victim - Far-end aggressor - Near-end aggressor	 A_v A_{fe} A_{ne}	 0.4 0.4 0.6	 V V V
Number of signal levels	L	2	-
Level separation mismatch ratio	R_{LM}	1	
Transmitter signal-to-noise ratio	SNR_{TX}	27	dB
Number of samples per unit interval	M	32	-
Decision feedback equalizer (DFE) length	N_b	5	UI
Normalized DFE coefficient magnitude limit, for n = 1 to N_b	$b_{max}(n)$	0.3	-
Random jitter, RMS	σ_{RJ}	0.01	UI
Dual-Dirac jitter, peak	A_{DD}	0.05	UI
One-sided noise spectral density	n_o	5.2×10^{-8}	V^2/GHz
Target detector error ratio	DER_0	10^{-15}	-

Table K. 6. *COM parameter values.*

Annex L. Chip-to-module 100Gb/s four-lane Attachment Unit Interface (CAUI-4)

This annex defines the functional and electrical characteristics for the optional chip-to-module 100 Gb/s four-lane Attachment Unit Interface (CAUI-4). **Figure L. 1** shows the relationship of the CAUI-4 chip-to-module interface to the ISO/IEC OSI reference model. The C2M interface provides the electrical characteristics and the associated compliance points, which can optionally be used when designing systems with pluggable module interfaces.



CAUI-4 = 100 Gb/s FOUR-LANE ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
RS-FEC = REED-SOLOMON FORWARD ERROR CORRECTION

Figure L. 1. *Example CAUI-4 chip-to-module relationship to the ISO/IEC OSI reference model and the IEEE 802.3 Ethernet model.*

The CAUI-4 link is comprised of a host CAUI-4 component, a CAUI-4 channel with associated insertion loss, and a module CAUI-4 component. **Figure L. 2** and **Equation L- 1** depict a typical CAUI-4 application and summarize the differential insertion loss budget associated with the chip-to-module application, which is shown in **Figure L. 3**.

The CAUI-4 C2M interface comprises independent data paths in each direction. Each data path contains four differential lanes, which are AC-coupled within the module. The nominal signaling rate for each lane is 25.78125GBd. The C2M interface is defined using a specification and a test methodology that is similar to that used for CEI-28G-VSR defined in OIF-CEI-03.1 [B54].

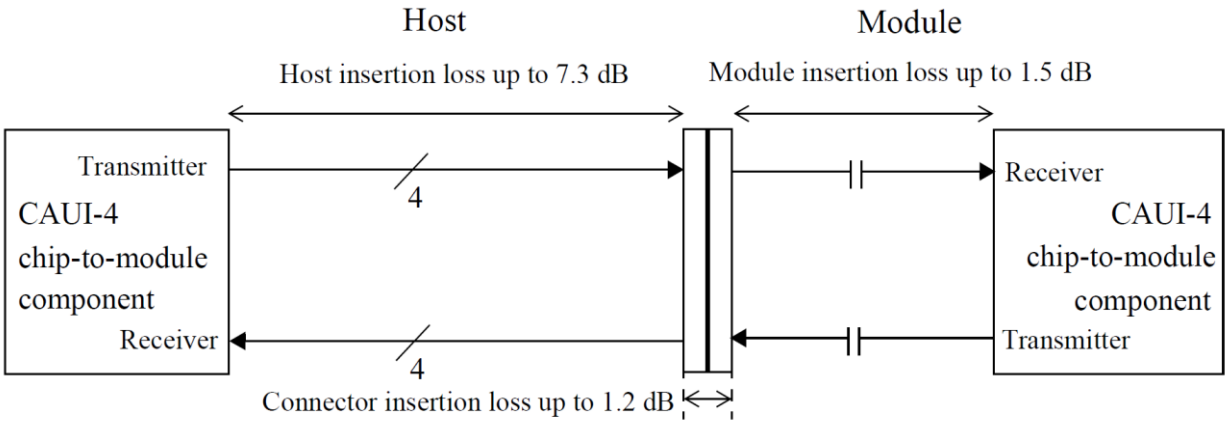


Figure L. 2. *Chip-to-module insertion loss budget at 12.89GHz.*

$$Insertion\_loss(f) \leq \begin{cases} 1.076 + 0.537\sqrt{f} + 0.566f & 0.01 \leq f < 14 \\ 1.076(-18 + 2f) & 14 \leq f \leq 18.75 \end{cases} \quad (dB)$$

Equation L- 1

where

Insertion_loss(f) is the CAUI-4 chip-to-module insertion loss

f is the frequency in GHz

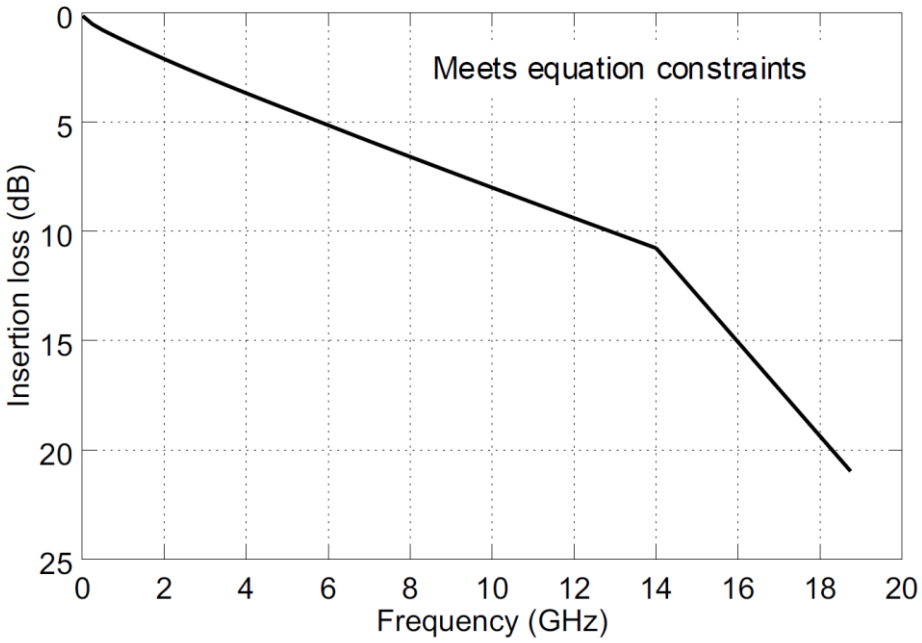


Figure L. 3. *CAUI-4 chip-to-module channel insertion loss.*

The BER shall be less than 10^{-15} with any errors sufficiently uncorrelated to ensure an acceptably high Mean Time To False Packet Acceptance (MTTFPA) assuming 64B/66B coding.

L.1. CAUI-4 chip-to-module compliance point definitions

The electrical characteristics for the CAUI-4 chip-to-module interface are defined at compliance points for the host and module, respectively. Reference test fixtures, called compliance boards, are used to access the electrical specification parameters. **Figure L. 4** depicts the location of compliance points when measuring host CAUI-4 compliance. The output of the HCB is used to verify the host electrical output signal at TP1a. Similarly, the input of the HCB at TP4a is used to verify the host input compliance.

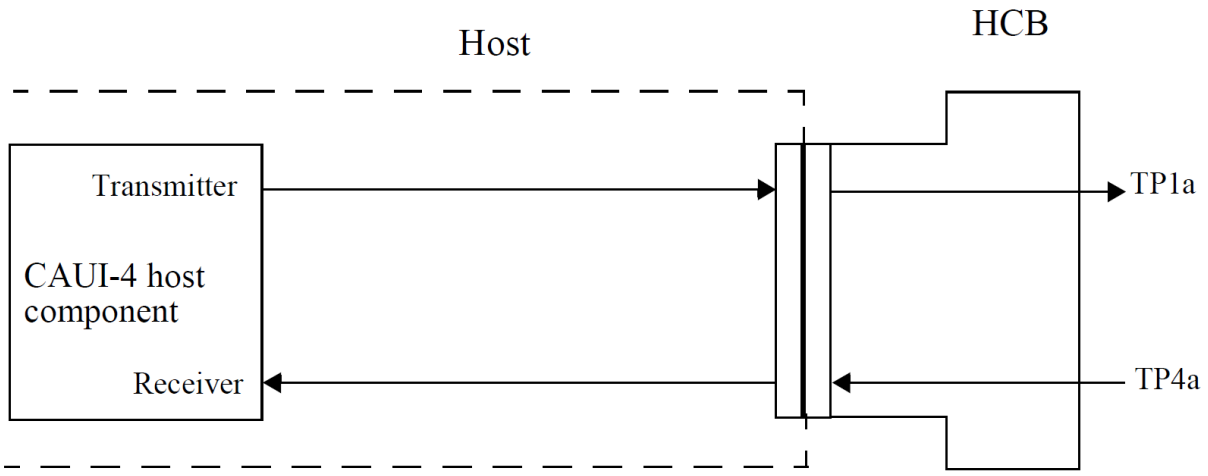


Figure L. 4. *Host CAUI-4 compliance points.*

Figure L. 5 depicts the location of compliance points when measuring module CAUI-4 compliance. The output of the MCB is used to verify the module electrical output signal at TP4. Similarly, the input of the MCB at TP1 is used to verify the module input compliance.

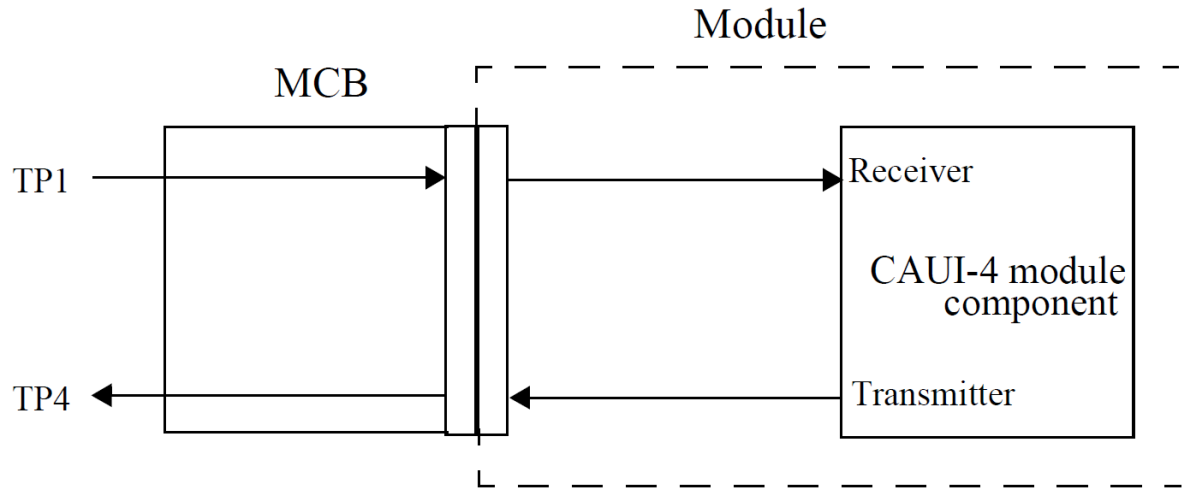


Figure L. 5. *Module CAUI-4 compliance points.*

L.2. CAUI-4 chip-to-module electrical characteristics

L.2.1. CAUI-4 host output characteristics

A CAUI-4 host output shall meet the specifications defined in **Table L. 1** if measured at TP1a.

Parameter	Value	Units
Signaling rate per lane (range)	25.78125 ± 100 ppm	GBd
DC common-mode output voltage (max)	2.8	V
DC common-mode output voltage (min)	-0.3	V
Single-ended output voltage (max)	3.3	V
Single-ended output voltage (min)	-0.4	V
AC common-mode output voltage (max, RMS)	17.5	mV
Differential peak-to-peak output voltage (max) - Transmitter disabled - Transmitter enabled	 35 900	mV
Eye width (min)	0.46	UI
Eye height A, differential (min)	95	mV
Eye height B, differential (min)	80	mV
Differential output return loss (min)	**Equation L- 2**	dB
Common to differential mode conversion return loss (min)	**Equation L- 3**	dB
Differential termination mismatch (max)	10	%
Transition time (min, 20% to 80%)	10	ps

Table L. 1. *CAUI-4 host output characteristics (at TP1a).*

A test system with a fourth-order Bessel-Thomson low-pass response with 33GHz 3dB bandwidth is to be used for all output signal measurements, unless otherwise specified.

L.2.1.1. Signaling rate and range

The CAUI-4 signaling rate is 25.78125GBd ± 100 ppm per lane. This translates to a nominal unit interval of 38.787879ps.

L.2.1.2. Signal levels

The differential output voltage v_{di} is defined to be the difference between the single-ended output voltages, $SL_i<p>$ minus $SL_i<n>$. The common-mode voltage v_{cmi} is defined to be one half of the sum of $SL_i<p>$ and $SL_i<n>$. These definitions are illustrated by **Figure L. 6**.

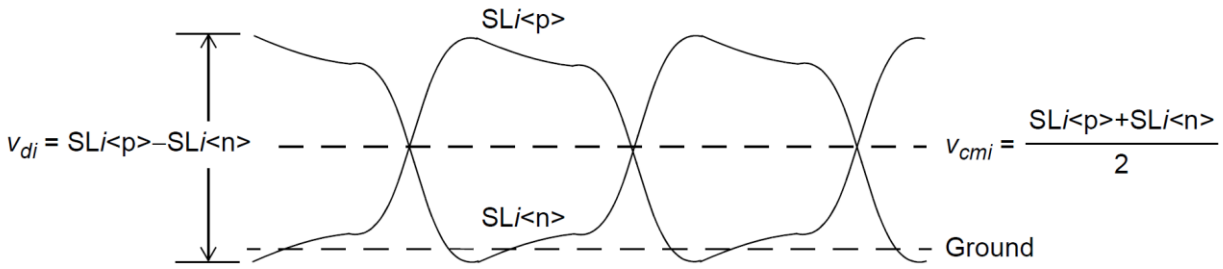


Figure L. 6. *Voltage definitions.*

The peak-to-peak differential output voltage is less than or equal to 900mV. The peak-to-peak differential output voltage is less than or equal to 35mV when the transmitter is disabled.

The DC common-mode output voltage and AC common-mode output voltage are defined with respect to signal ground.

L.2.1.3. Output return loss

The differential output return loss, in dB, of the output is shown in **Equation L- 2** and illustrated in **Figure L. 7**. This output requirement applies to all valid output levels. The reference impedance for differential return loss measurements is 100Ω.

$$RLd(f) \geq \begin{Bmatrix} 9.5 - 0.37f & 0.01 \leq f < 8 \\ 4.45 - 7.4log_{10}\left(\frac{f}{14}\right) & 8 \leq f \leq 19 \end{Bmatrix} (dB)$$

Equation L- 2

where

RLd is the CAUI-4 chip-to-module host output differential return loss

f is the frequency in GHz

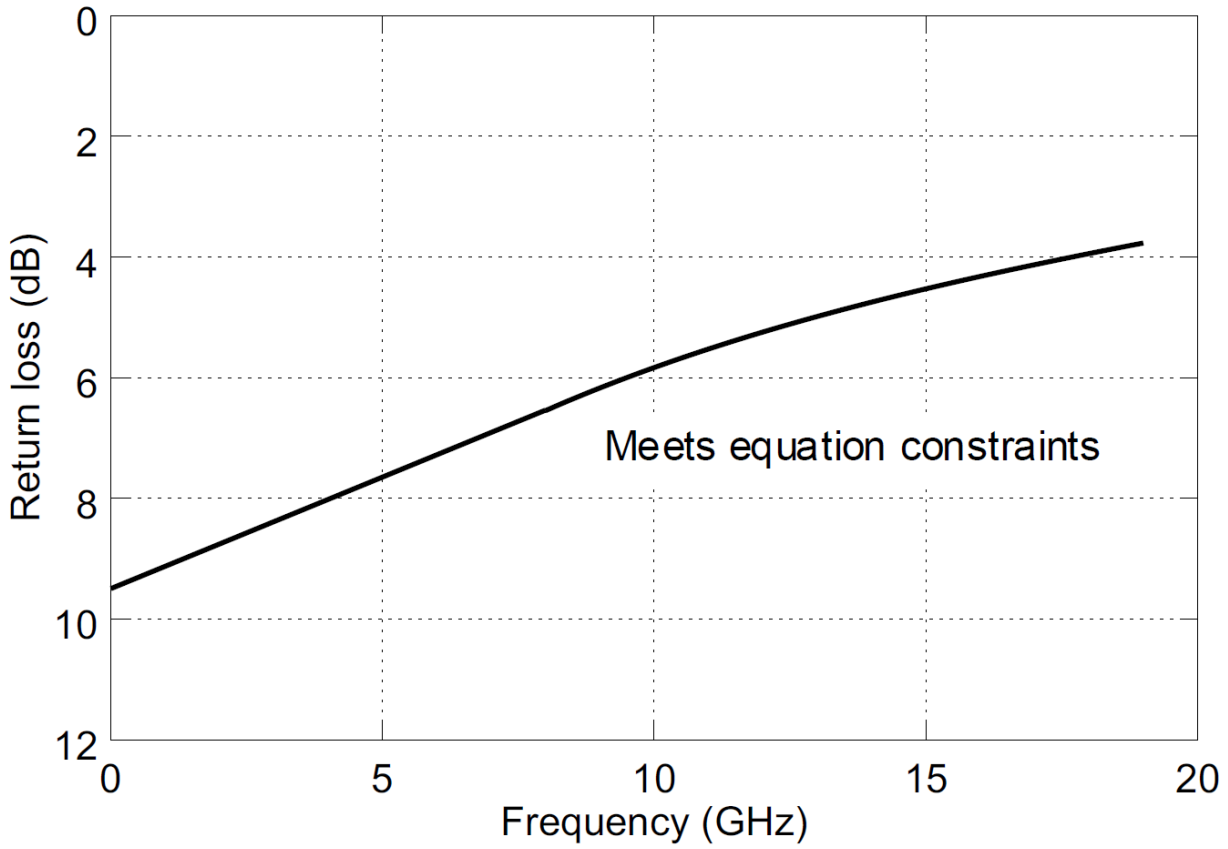


Figure L. 7. *Output differential return loss.*

Common to differential output conversion return loss, in dB, of the output is shown in **Equation L- 3** and illustrated in **Figure L. 8**.

$$RLdc(f) \geq \begin{Bmatrix} 22 - 20\left(\frac{f}{25.78}\right) & 0.01 \leq f < 12.89 \\ 15 - 6\left(\frac{f}{25.78}\right) & 12.89 \leq f \leq 19 \end{Bmatrix} \ (dB)$$

Equation L- 3

where

RLdc is the CAUI-4 chip-to-module output common to differential mode conversion return loss

f is the frequency in GHz

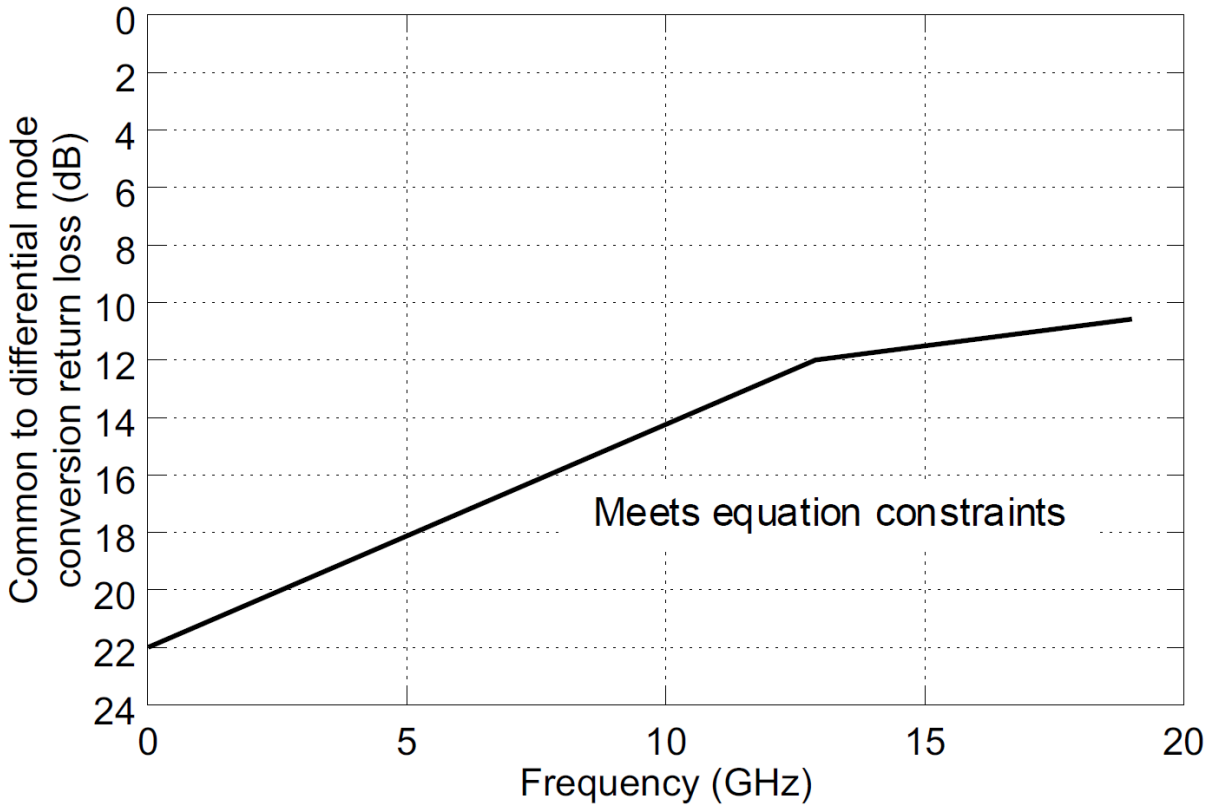


Figure L. 8. *Output common to differential mode conversion return loss.*

L.2.1.4. Host output eye width and eye height

Figure L. 9 shows an example host output eye width and eye height test configuration. Host output eye width and eye height are measured at TP1a using compliance boards defined in this annex. The host output eye is measured using a reference receiver with a CTLE defined below. The recommended CTLE peaking value is used for the host output eye measurements. In addition, it is provided to the module via the variable *Recommended_CTLE_value*. If an MDIO is implemented, this variable is accessible in the module through register 1.179. Eye width and eye height measurement methodology is described below. All counter-propagating signals shall be asynchronous to the copropagating signals using Pattern 5 (with or without FEC encoding), Pattern 3, or a valid 100GBASE-R signal. Patterns 3 and 5 are described in **Table 2-16**. For the case where Pattern 3 is used with a common clock, there is at least 31 UI delay between the PRBS31 patterns on one lane and any other lane. The crosstalk generator is calibrated

at TP4 with target differential peak-to-peak amplitude of 900mV and target transition time of 12ps.

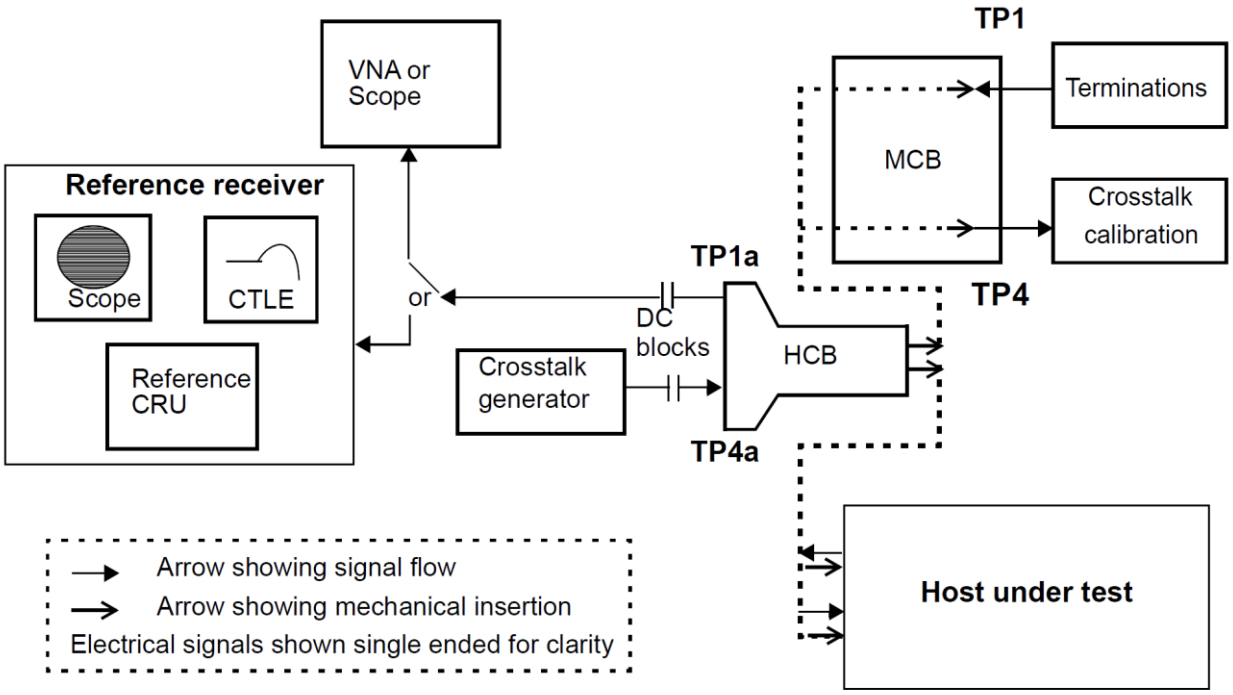


Figure L. 9. *Example host output test configuration.*

L.2.1.5. Reference receiver for host output eye width and eye height evaluation

The reference receiver is used to measure host eye width and eye height. The reference receiver includes a selectable CTLE, which is described by **Equation L- 4** with coefficients given in **Table L. 2** and illustrated in **Figure L. 10**. The equalizer may be implemented in software; however, the measured signal is not averaged.

$$H(f) = \frac{GP_1P_2}{Z_1} \times \frac{j2\pi f + Z_1}{(j2\pi f + P_1)(j2\pi f + P_2)}$$

Equation L- 4

where

H(f) is the CTLE transfer function

G is the CTLE gain

P_1, P_2 are the CTLE poles in Grad/s

Z_1 is the CTLE zero in Grad/s

j is the square root of –1

f is the frequency in GHz

Peaking (dB)	G	$\frac{P_1}{2\pi}$	$\frac{P_2}{2\pi}$	$\frac{Z_1}{2\pi}$
1	0.89125	18.6	14.1	8.364
2	0.79433	18.6	14.1	7.099
3	0.70795	15.6	14.1	5.676
4	0.63096	15.6	14.1	4.9601
5	0.56234	15.6	14.1	4.358
6	0.50119	15.6	14.1	3.844
7	0.44668	15.6	14.1	3.399
8	0.39811	15.6	14.1	3.012
9	0.35481	15.6	14.1	2.672

Table L. 2. *Reference CTLE coefficients.*

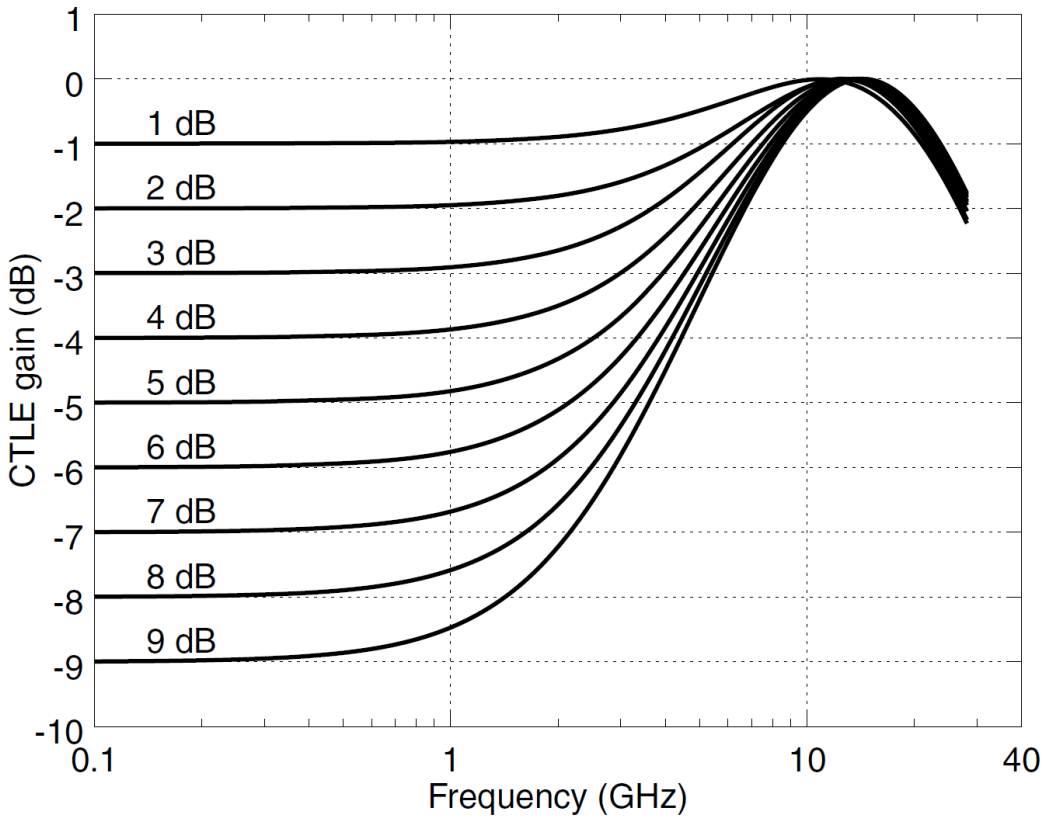


Figure L. 10. *Selectable continuous time linear equalizer (CTLE) characteristic.*

L.2.2. CAUI-4 module output characteristics

A CAUI-4 module output shall meet the specifications defined in **Table L. 3** if measured at TP4. A test system with a fourth-order Bessel-Thomson low-pass response with 33GHz 3dB bandwidth is to be used for all output signal measurements, unless otherwise specified.

Parameter	Value	Units
Signaling rate per lane (range)	25.78125 ± 100 ppm	GBd
AC common-mode output voltage (max, RMS)	17.5	mV
Differential output voltage (max)	900	mV
Eye width (min)	0.57	UI
Eye height, differential (min)	228	mV
Vertical eye closure (max)	5.5	dB
Differential output return loss (min)	**Equation L- 2**	dB
Common to differential mode conversion return loss (min)	**Equation L- 3**	dB
Differential termination mismatch (max)	10	%
Transition time (min, 20% to 80%)	12	ps
DC common mode voltage (min) [a]	-350	mV
DC common mode voltage (max) [a]	2850	mV
[a] DC common mode voltage is generated by the host. Specification includes effects of ground offset voltage.		

Table L. 3. *CAUI-4 module output characteristics (at TP4).*

L.2.2.1. Module output eye width and eye height

Module output eye width is greater than 0.57 UI. Module output eye height is greater than 228mV. **Figure L. 11** depicts an example module output eye width and eye height test configuration. Module output eye width and eye height are measured at TP4 using compliance boards. The module output eye is measured using a reference receiver with a CTLE. All counterpropagating signals shall be asynchronous to the co-propagating signals using Pattern 5 (with or without FEC encoding), Pattern 3, or a valid 100GBASE-R signal. Patterns 3 and 5 are described in **Table 2-16**. For the case where Pattern 3 is used with a common clock, there is at least 31 UI delay between the PRBS31 patterns

on one lane and any other lane. The crosstalk generator is calibrated at TP1a with target differential peak-to-peak amplitude of 900mV and target transition time of 19ps.

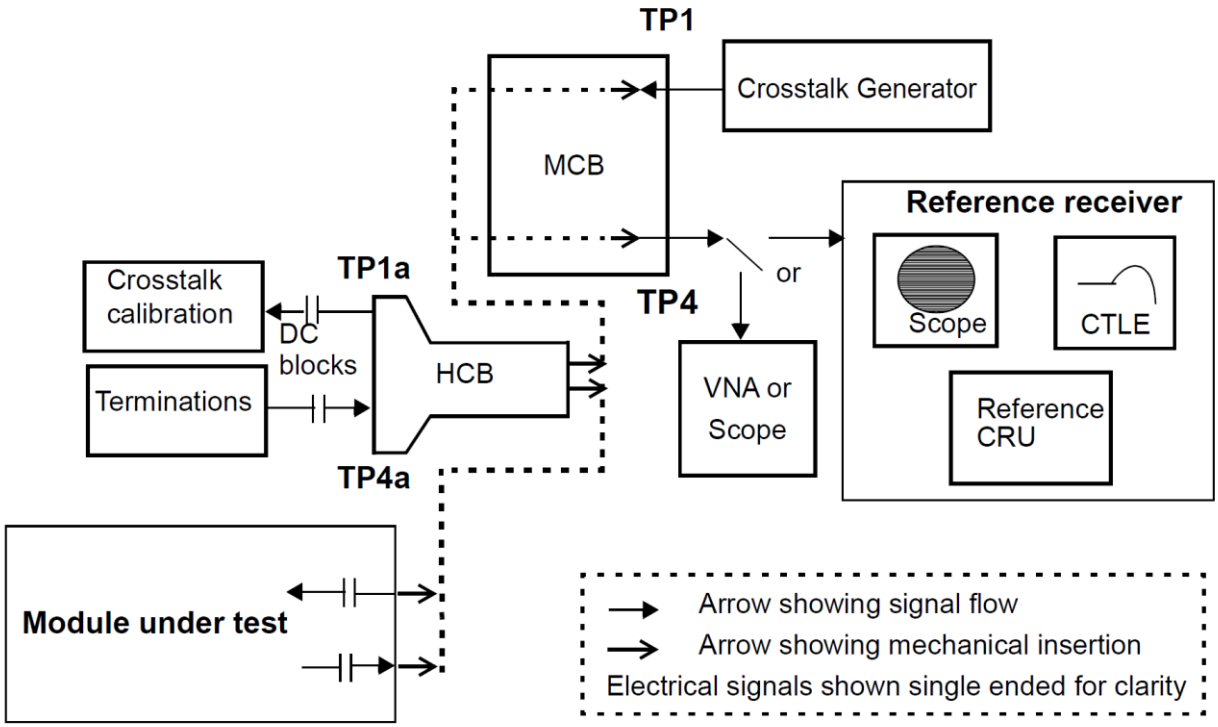


Figure L. 11. *Example module output test configuration.*

Annex M. Parallel Physical Interface (nPPI) for 100GBASE-SR10 (CPPI)

The Parallel Physical Interface (nPPI) is an optional instantiation of the PMD service interface for the PMDs in Section **2.6**, which is described in Sections **2.6** and **2.2**. It allows the construction of compact optical transceiver modules for 100GBASE-SR10 with no clock and data recovery circuits inside. The 100 Gb/s Parallel Physical Interface (CPPI, ten lanes) with 100GBASE-SR10, where “n” replaced by “C”.

The PMD and PMA attached to the nPPI are required to comply with the delay, Skew, and Skew Variation requirements in Sections **2.6**, and **2.5** as appropriate. The PMD MDIO function mapping given in Section **2.6** may apply. The PMD functional specifications are as given in Section **2.4**.

This annex is arranged as follows: following the overview, a reminder of the block diagram, a brief introduction to the test points, and lane assignments. This annex contains the electrical specifications for nPPI from host to module (Tx side) and then module to host (Rx side). Test points, compliance boards, and electrical parameters are defined below. This annex is very closely related to Section **2.6**.

M.1. Block diagram and test points

The PMD block diagram is shown in **Figure 2-15**. **Figure 2-14** shows the test points.

The nPPI is standardized at the test points described in Section **2.6.4**. The transmit side electrical signal (host electrical output and module electrical input) is defined at the output of the Host Compliance Board (TP1a), and other specifications of the module electrical input port are defined at the input of the Module Compliance Board (TP1). The receive side electrical signal (module electrical output and host electrical input) is defined at the output of the Module Compliance Board (TP4), and other specifications of the host electrical input port are defined at the input of the Host Compliance Board (TP4a). Test points and compliance boards are defined more thoroughly below.

M.2. Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for CPPI. While it is expected that a PMD will map electrical lane i to optical lane *i* and vice versa,

there is no need to define the physical ordering of the lanes, as the PCS is capable of receiving the lanes in any arrangement.

M.3. Electrical specifications for nPPI

The signaling rate for a lane of an CPPI interface shall be as defined in **Table M. 2**. The Sections below specify the host to module (Tx side) and module to host (Rx side) respectively of the nPPI. Parameters are also defined below. A recommended PCB ("channel") response for the host (PMA) is provided below. Test points are defined also below.

Parameter description	Min	Max	Units	Conditions
Single ended output voltage	-0.3	4	V	Referred to signal common
AC common-mode output voltage	-	15	mV	RMS
Termination mismatch at 1MHz	-	5	%	
Differential output return loss	See this Annex	-	dB	
Output transition time, 20% to 80%	28	-	ps	
J2 Jitter output	-	0.17	UI	
J9 Jitter output	-	0.29	UI	
Data Dependent Pulse Width Shrinkage (DDPWS)	-	0.07	UI	
Q_{sq} for CPPI	43	-	V/V	
	Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	0.11, 0.31 95, 350		UI mV	Hit ratio = 5×10^{-5}
Crosstalk source VMA, each input lane	700		mV	At TP4
Crosstalk source transition times, 20% to 80%	34		ps	At TP4

Table M. 1. *nPPI host electrical output specifications at TP1a.*

Parameter description	Test point	Min	Max	Units	Conditions
Single ended input voltage tolerance [a]	TP1a	-0.3	4	V	Referred to TP1 signal common
AC common-mode input voltage tolerance	TP1a	15	-	mV	RMS
Differential input return loss	TP1	See this Annex	-	dB	10MHz to 11.1GHz
Differential to common-mode input return loss	TP1	10	-	dB	10MHz to 11.1GHz
J2 Jitter tolerance	TP1a	0.17	-	UI	
J9 Jitter tolerance	TP1a	0.29	-	UI	
Data Dependent Pulse Width Shrinkage (DDPWS) tolerance	TP1a	0.07	-	UI	
		Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	TP1a	0.11, 0.31 95, 350		UI mV	Hit ratio = 5×10^{-5}
Crosstalk calibration signal VMA	TP4	850		mV	While calibrating compliance signal [b]
Crosstalk calibration signal transition times, 20% to 80%	TP4	34		ps	

[a] The single ended input voltage tolerance is the allowable range of the instantaneous input signals.

[b] The crosstalk calibration signals are applied to the mated HCB-MCB at TP4a and measured at TP4 following the same principles as the host electrical input calibration. They are removed before testing.

Table M. 2. *nPPI module electrical input specifications at TP1 and TP1a.*

M.3.1. nPPI host to module electrical specifications

Each output lane and signal of the nPPI host (PMA), if measured at TP1a with the specified crosstalk signals applied on all input lanes, shall meet the specifications of **Table M. 1** per the definitions of this annex. Each lane of the nPPI module (PMD) electrical input, if measured at TP1 and TP1a with all Rx lanes (module output) operating,

shall meet the specifications of **Table M. 2** per the definitions of this annex. The module electrical input shall be AC-coupled, i.e., it shall present a high DC common-mode impedance at TP1.

M3.1.1. Differential return losses at TP1 and TP1a

From 10MHz to 11.1GHz, the magnitude in decibels of the module differential input return loss at TP1 and the host differential output return loss at TP1a shall not exceed the limit given in **Equation M- 1** and illustrated in **Figure M. 1**.

$$Return\_loss(f) \geq \begin{cases} 12 - 2\sqrt{f} & 0.01 \leq f < 4.11 \\ 6.3 - 13log_{10}\left(\frac{f}{5.5}\right) & 4.11 \leq f \leq 11.1 \end{cases} (dB)$$

Equation M- 1

where

Return_loss(f) is the return loss at frequency *f*

f is the frequency in GHz

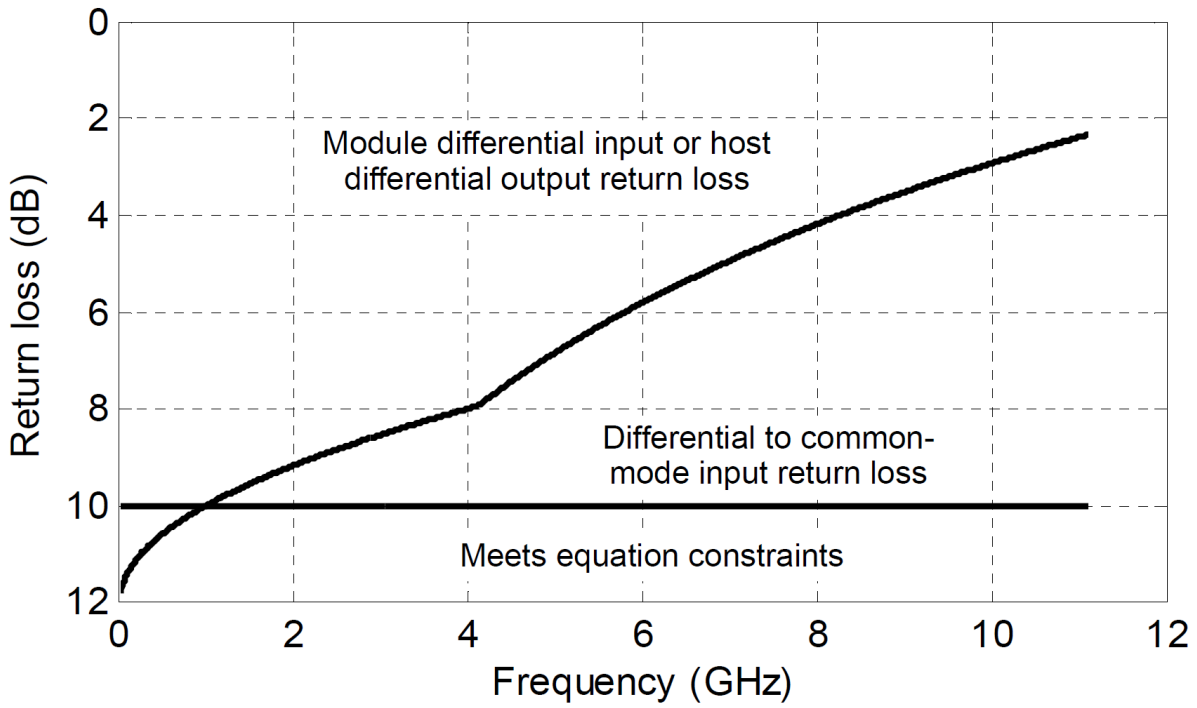


Figure M. 1. *Return loss specifications.*

M.3.2. nPPI module to host electrical specifications

Each electrical output lane and signal of the nPPI module (PMD), if measured at TP4 shall meet the specifications of **Table M. 3** per the definitions of this annex while the specified crosstalk sources are applied to the module's electrical input. Each lane of the nPPI host (PMA) input shall meet the specifications of **Table M. 4** at TP4 and/or TP4a per the definitions of this annex. The module electrical output shall be AC-coupled, i.e., it shall present a high DC common-mode impedance at TP4. There may be various methods for AC-coupling in actual implementations.

Parameter description	Min	Max	Units	Conditions
Single ended output voltage tolerance	-0.3	4	V	Referred to signal common
AC common-mode output voltage (RMS)	-	7.5	mV	
Termination mismatch at 1MHz	-	5	%	
Differential output return loss	See this Annex	-	dB	10MHz to 11.1GHz
Output transition time, 20% to 80%	28	-	ps	
J2 Jitter output	-	0.42	UI	
J9 Jitter output	-	0.65	UI	
	Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	0.29, 0.5 150, 425		UI mV	Hit ratio = 5×10^{-5}
Crosstalk source VMA, each lane	700		mV	At TP1a
Crosstalk source transition times, 20% to 80%	37		ps	At TP1a

Table M. 3. *nPPI module electrical output specifications at TP4.*

Parameter description	Test point	Min	Max	Units	Conditions
Single ended input voltage [a]	TP4	-0.3	4	V	Referred to signal common
AC common-mode input voltage tolerance	TP4	7.5	-	mV	RMS

Differential input return loss	TP4a	See this Annex	-	dB	
Differential to common-mode input return loss	TP4a	10	-	dB	10MHz to 11.1GHz
Host input signal tolerance, interface BER limit		-	10^{-12}	-	
Conditions of host electrical receiver signal tolerance test: [b]					
		Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	TP4	0.29, 0.5 150, 425		UI mV	Hit ratio = 5×10^{-5}
Transition time, 20% to 80%		34		ps	
J2 Jitter	TP4	0.42		UI	
J9 Jitter	TP4	0.65		UI	
Data Dependent Pulse Width Shrinkage (DDPWS)		0.34		UI	
VMA of aggressor lanes	TP4	850		mV	
Crosstalk calibration signal VMA	TP1a	700		mV	
Crosstalk calibration signal transition times, 20% to 80%	TP1a	37		ps	
[a] The host is required to tolerate (work correctly with) input signals with instantaneous voltages anywhere in the specified range. [b] The specification values are test conditions for measuring signal tolerance and are not characteristics of the host.					

Table M. 4. *nPPI host electrical input specifications at TP4 and TP4a.*

M.3.2.1. Differential return losses at TP4 and TP4a

From 10MHz to 11.1GHz, the magnitude in decibels of the module differential output return loss at TP4 and the host differential input return loss at TP4a shall not exceed the limit given in **Equation M- 2** and illustrated in **Figure M. 1**.

$$Return\_loss(f) \geq \begin{cases} 12 - 2\sqrt{f} & 0.01 \leq f < 4.11 \\ 6.3 - 13log_{10}\left(\frac{f}{5.5}\right) & 4.11 \leq f \leq 11.1 \end{cases} \ (dB)$$

Equation M- 2

where

Return_loss(f) is the return loss at frequency *f*

f is the frequency in GHz

M.4. Definitions of electrical parameters and measurement methods

Test points, compliance boards, test patterns and parameters are defined in this annex and Section **2.6.4**. Multi-lane testing considerations are given in Section **2.6.4**.

M.4.1. Test points and compliance boards

Figure M. 2 shows the six test points for 100GBASE-SR10. These are TP1, TP1a, TP2, TP3, TP4, and TP4a; four of these are Skew points SP2, SP3, SP4, and SP5 as shown. **Figure M. 2** also shows the substitution of compliance boards for module (PMD) or host (PMA). These compliance boards are defined to connect generic test equipment to the module and host for test purposes. The module can be plugged into the MCB, which has specified electrical parameters. The HCB, which also has specified electrical parameters, can plug into the host. The MCB and the HCB can be plugged together for calibration of compliance signals and to check the electrical parameters of the boards. **Table M. 5** shows the parameters or signals measured at each point. Also, TP0 and TP5 define the host ends of the electrical channel, at the PMA IC.

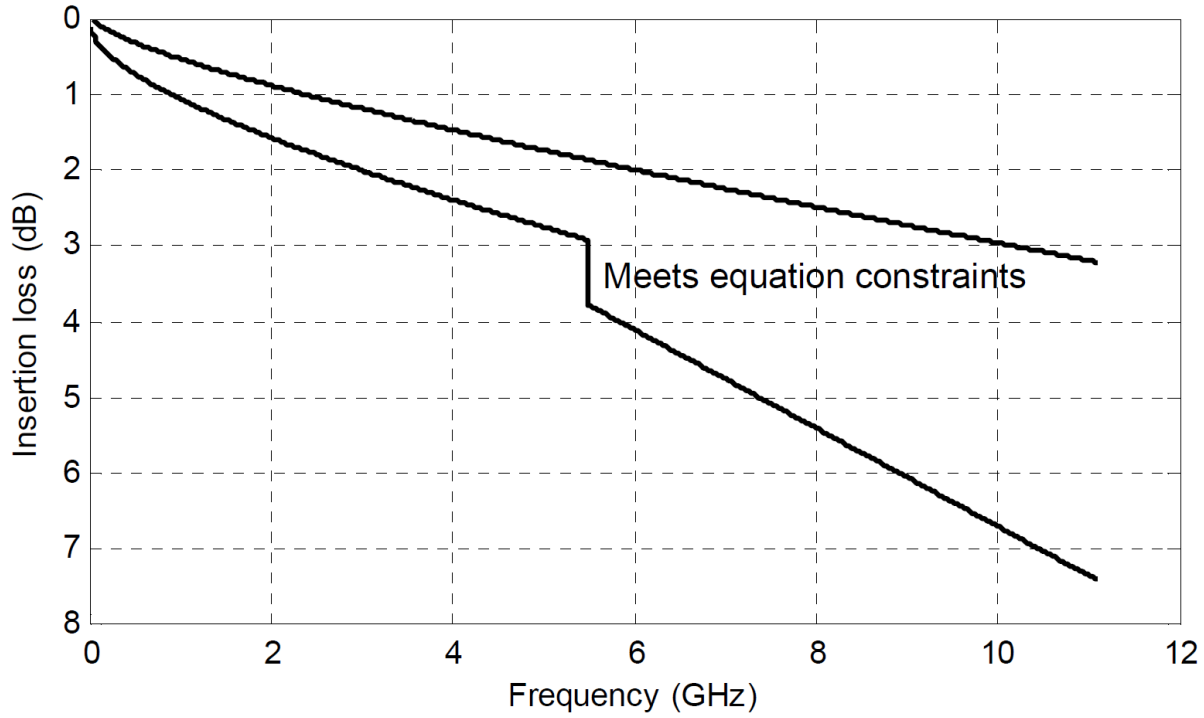


Figure M. 2. *Differential insertion loss limits of mated HCB-MCB.*

Test point	Direction	Parameter
TP1	Looking downstream into module transmitter input	Module transmitter input return loss
TP1a	Looking upstream into host transmitter output	Host transmitter output signal and output return loss, module transmitter compliance signal calibration, host receiver compliance crosstalk signal calibration
TP4	Looking upstream into module receiver output	Module receiver output signal and output return loss, host receiver compliance signal calibration
TP4a	Looking downstream into host receiver input	Host receiver input return loss

Table M. 5. *Parameters defined at each test point.*

M.4.1.1. Compliance board parameters

The electrical characteristics of the HCB and MCB are given in this annex. If boards are used that do not match the specifications given, the measurement results for nPPI

shall be corrected for the differences between the actual HCB or MCB's properties and the reference differential insertion losses given in this annex. As it may be impractical to correct eye measurements for a board with differential insertion loss outside the limits given in this annex, such boards shall not be used. Boards that do not meet the specifications for mated HCB-MCB in this annex shall not be used.

M.4.1.2. Reference insertion losses of HCB and MCB

The reference differential insertion loss in decibels of the HCB and of the MCB, excluding the module connector, are given in **Equation M- 3** and **Equation M- 4** respectively, and illustrated in **Figure M. 2**.

For the HCB,

$$Insertion\_loss(f) = 0.01 + 0.3\sqrt{f} + 0.11f \quad 0.01 \le f < 11.1$$

Equation M- 3

where

Insertion_loss(f) is the insertion loss at frequency *f*

f is the frequency in GHz

For the MCB,

$$Insertion\_loss(f) = 0.0006 + 0.16\sqrt{f} + 0.0587f \quad 0.01 \le f < 11.1$$

Equation M- 4

where

Insertion_loss(f) is the insertion loss at frequency *f*

f is the frequency in GHz

M.4.1.3. Electrical specifications of mated HCB and MCB

The limits on the differential insertion loss in decibels of the mated HCB and MCB (in either direction) are given in **Equation M- 5** (which defines the minimum insertion loss) and **Equation M- 6** (which defines the maximum insertion loss). These limits are illustrated in **Figure M. 3**.

$$Insertion_{loss(f)} \geq -0.11 + 0.46\sqrt{f} + 0.167f \qquad 0.01 \leq f < 11.1$$

Equation M- 5

$$Insertion\_loss(f) \leq \begin{Bmatrix} 0.029 + 0.861\sqrt{f} + 0.158f & 0.01 \leq f < 5.5 \\ 0.2 + 0.65f & 5.5 \leq f \leq 11.1 \end{Bmatrix} \ (dB)$$

Equation M- 6

where

Insertion_loss(f) is the insertion loss at frequency *f*

f is the frequency in GHz

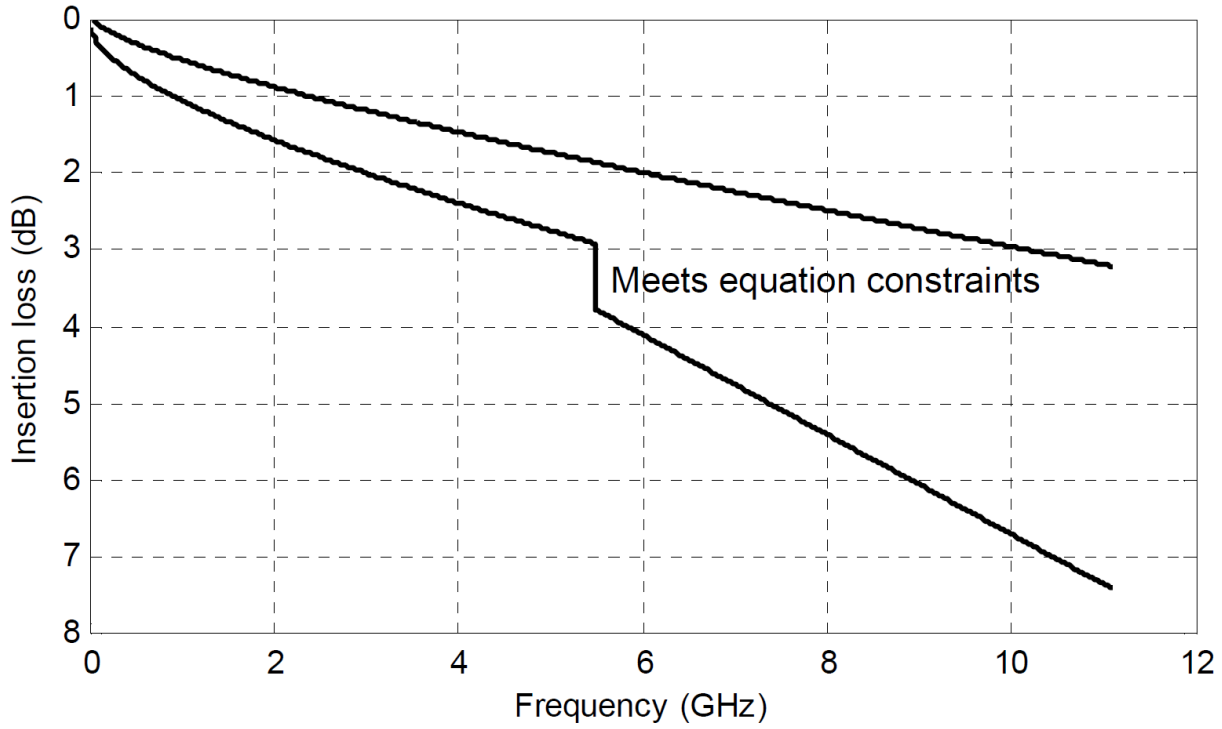


Figure M. 3. *Differential insertion loss limits of mated HCB-MCB.*

The limit on the differential return loss in decibels of the mated HCB and MCB, looking into the HCB or looking into the MCB, is given in **Equation M- 7**, and illustrated in **Figure M. 4**.

$$Return\_loss(f) \geq \begin{cases} 20 - 2f & 0.01 \leq f < 2.5 \\ 15 & 2.5 \leq f < 5 \\ 13.8 - 28.85log_{10}\left(\frac{f}{5.5}\right) & 5 \leq f \leq 11.1 \end{cases} (dB)$$

Equation M- 7

where

Return_loss(f) is the return loss at frequency *f*

f is the frequency in GHz

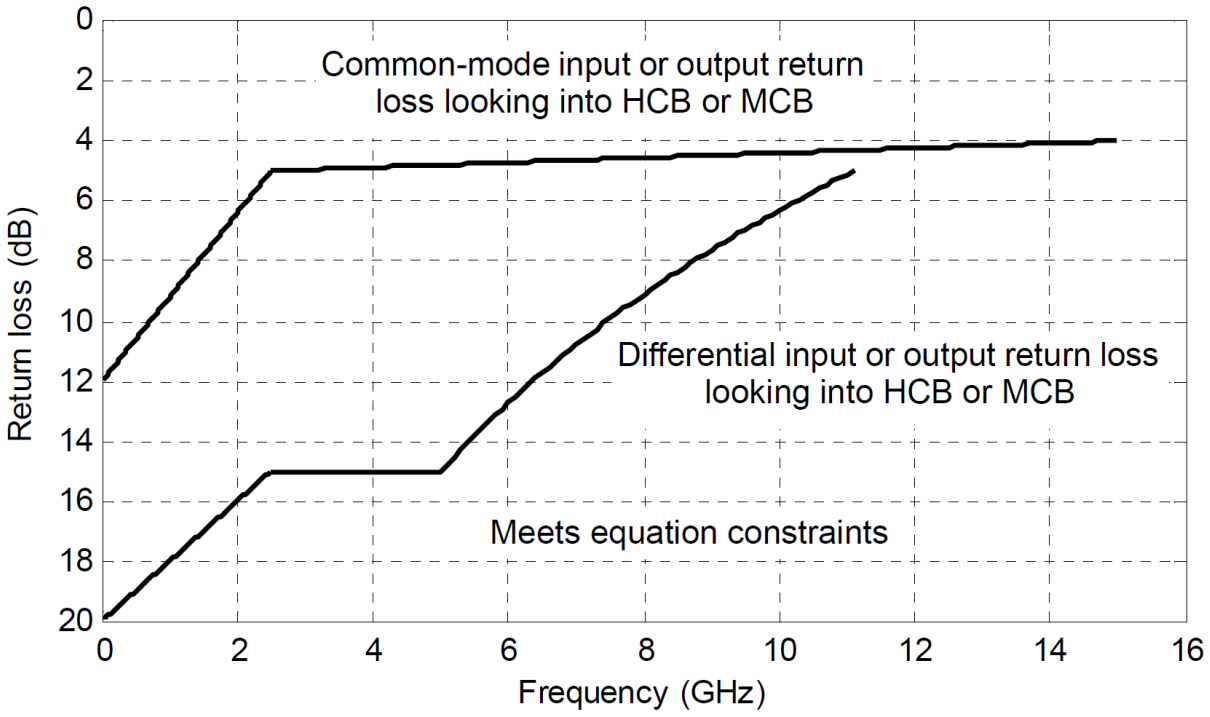


Figure M. 4. *Return loss limits of mated HCB-MCB.*

The limit on the common-mode return loss in decibels of the mated HCB and MCB is given in **Equation M- 8** and illustrated in **Figure M. 4**.

$$Return\_loss(f) \geq \begin{cases} 12 - 2.8f & 0.01 \leq f < 2.5 \\ 5.2 - 0.08f & 2.5 \leq f \leq 15 \end{cases} (dB)$$

Equation M- 8

where

Return_loss(f) is the return loss at frequency *f*

f is the frequency in GHz

The limit on the differential to common-mode conversion loss in decibels of the mated HCB and MCB, for input to HCB and output from MCB, or input to MCB and output from HCB, is given in **Equation M- 9** and illustrated in **Figure M. 5**.

$$Mode\_conversion\_loss(f) \geq \begin{cases} 30 - 2.91f & 0.01 \leq f < 5.5 \\ 14 & 5.5 \leq f \leq 15 \end{cases} \ (dB)$$

Equation M- 9

where

Mode_conversion_loss(f) is the mode conversion loss at frequency *f*

f is the frequency in GHz

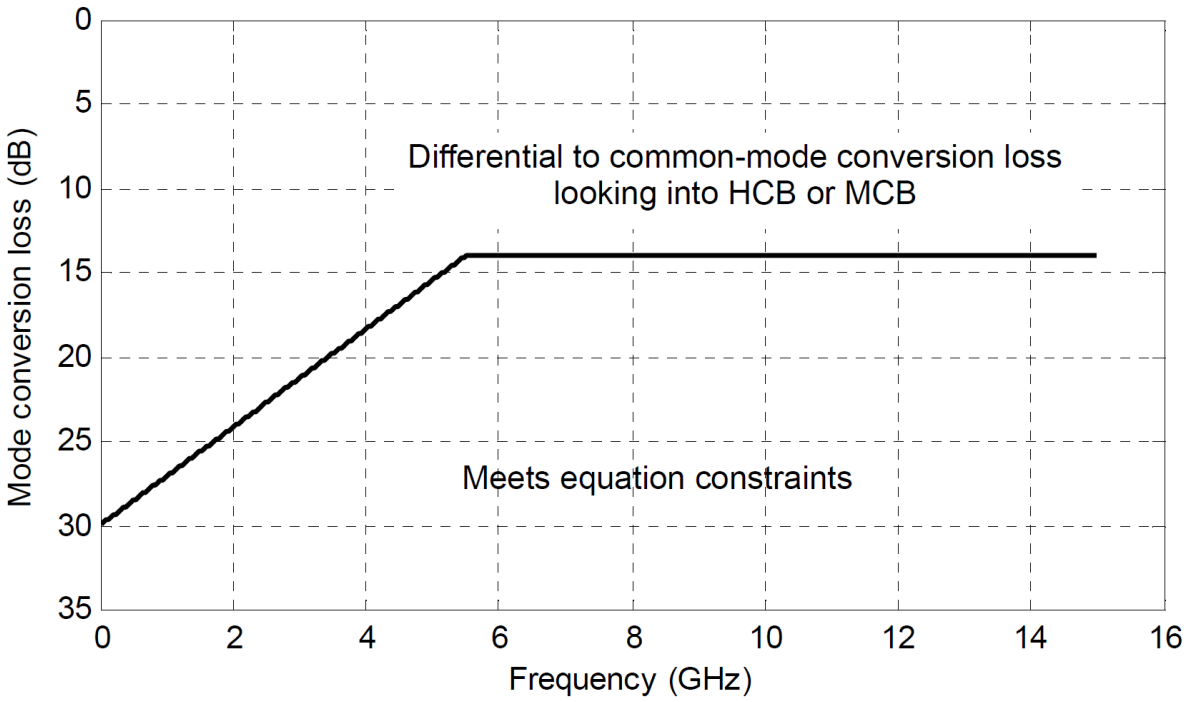


Figure M. 5. *Differential to common-mode conversion loss limit of mated HCB-MCB.*

M.4.2. Test patterns and related subsections

While compliance is to be achieved in normal operation, specific test patterns are defined for measurement consistency and to enable measurement of some parameters. **Table 2-16** lists the defined test patterns, and **Table M. 6** gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the

subclauses in which each parameter is defined. Multi-lane testing considerations are given in Section **2.6.4**. As Pattern 3 is more demanding than Pattern 5 (which itself is the same or more demanding than other 100GBASE-R bit streams) an item which is compliant using Pattern 5 is considered compliant even if it does not meet the required limit using Pattern 3.

Parameter	Pattern	Related subsection
J2 Jitter	3, 5, or valid 100GBASE-SR10 signal	**2.6.4**
J9 Jitter	3 or 5	**2.6.4**
Data Dependent Pulse Width Shrinkage (DDPWS)	4	**Annex M**
AC common-mode voltage	3, 5, or valid 100GBASE-SR10 signal	**Annex M**
Transition time	Square wave or 4	**Annex M**
Electrical waveform (eye mask)	3, 5, or valid 100GBASE-SR10 signal	**2.6.4, Annex M**
Q_{sq}	Square wave or 4	**Annex M**
Host electrical receiver signal tolerance	3 or 5	**Annex M**

Table M. 6. *Test patterns and related subclauses.*

M.4.3. Parameter definitions

In addition to the parameter definitions in the following subclause, some definitions with dual use (both optical and electrical) are given in Section **2.6.4**.

M.4.3.1. AC common-mode voltage

The common-mode voltage of a differential signal at any time is the average of signal+ and signal– at that time. RMS AC common-mode voltage may be calculated by applying the histogram function over 1 UI to the common-mode signal. As AC common-mode generation is very sensitive to the cable or oscilloscope delay mismatch, it is recommended to delay match the oscilloscope inputs for any measurements.

M.4.3.2. Termination mismatch

Termination mismatch is the percentage difference between the two low-frequency impedances to common of a differential electrical port. Termination mismatch is defined as shown in **Equation M- 10**.

$$\Delta Z_M = 2 \times \frac{|Z_p - Z_n|}{Z_p + Z_n} \times 100 \ \%$$

Equation M- 10

Termination mismatch can be measured by applying a low-frequency test tone to the differential inputs as shown in **Figure M. 6**. The test frequency must be high enough to overcome the high pass effects of any AC-coupling capacitors. The measured differential output or input impedance is designated by Z_{diff}.

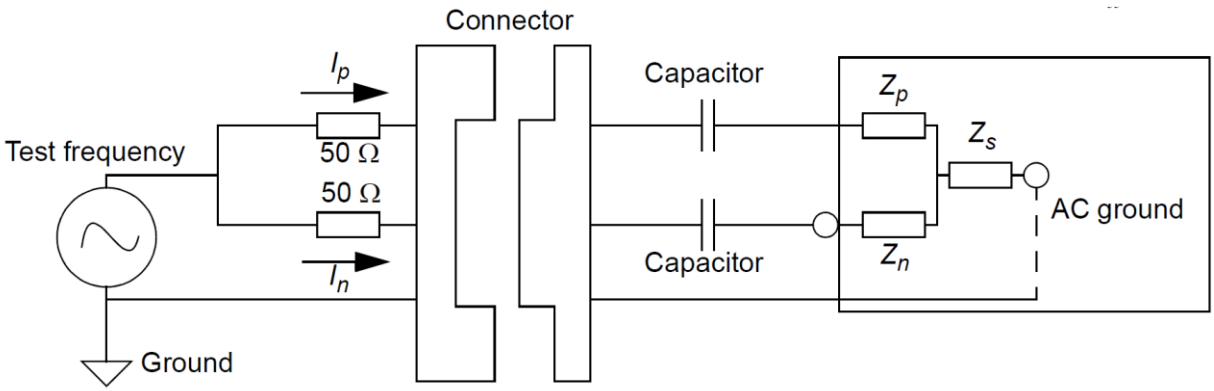


Figure M. 6. *Measurement of AC termination mismatch.*

For a 100Ω port, low-frequency termination mismatch is then given by **Equation M- 11**.

$$\Delta Z_M = 2 \times \frac{|I_p - I_n|}{I_p + I_n} \times \frac{Z_{diff} + 100}{Z_{diff}} \times 100 \quad \%$$

Equation M- 11

where I_p and I_n are the currents flowing into the port as shown in **Figure M. 6**.

Z_s is the effective series impedance between the terminations Z_p and Z_n and the AC ground.

M.4.3.3. Transition time

In this annex, transition times (rise and fall times) are defined as the time between the 20% and 80% times, or 80% and 20% times, respectively, of isolated edges.

If the test pattern is the square wave with eight ones and eight zeros, the 0% level and the 100% level are as defined by the OMA measurement procedure.

If the test pattern is PRBS9, the transitions within sequences of five zeros and four ones, and nine ones and five zeros, respectively, are measured. These are bits 10 to 18 and 1 to 14, respectively, where bits 1 to 9 are the run of nine ones. In this case, the 0% level and the 100% level may be estimated as *ZeroLevel* and *ZeroLevel* + *MeasuredOMA* in the TWDP code, or by the average signal within windows from –3 UI to –2 UI and from 2 UI to 3 UI relative to the edge.

For electrical signals, the waveform is observed through a 12GHz low-pass filter response (such as a Bessel-Thomson response).

NOTE: This definition is not the same as the rise and fall times typically reported by an oscilloscope from an eye diagram, which take all the edges into account.

M.4.3.4. Data Dependent Pulse Width Shrinkage (DDPWS)

An oscilloscope, time interval analyzer, or other instrument with equivalent capability may be used to measure Data Dependent Pulse Width Shrinkage (DDPWS). A CRU as defined in Section **2.6.4** is used to trigger the oscilloscope. A repeating PRBS9 pseudo-random test pattern, 511 bits long, is used. For electrical jitter measurements, the measurement bandwidth is 12GHz (such as a Bessel-Thomson response). If the measurement bandwidth affects the result, it can be corrected for by post-processing. However, a bandwidth above 12GHz is expected to have little effect on the results.

The crossing level is the average value of the entire waveform being measured. The instrument is synchronized to the pattern repetition frequency and the waveforms, or the crossing times are averaged sufficiently to remove the effects of random jitter and noise in the system. The PRBS9 pattern has 128 positive-going transitions and 128 negative-going transitions. The crossing times t_1 to t_{256} of each transition of the averaged waveform (when the averaged waveform crosses the crossing level) are found as shown in **Figure M. 7**.

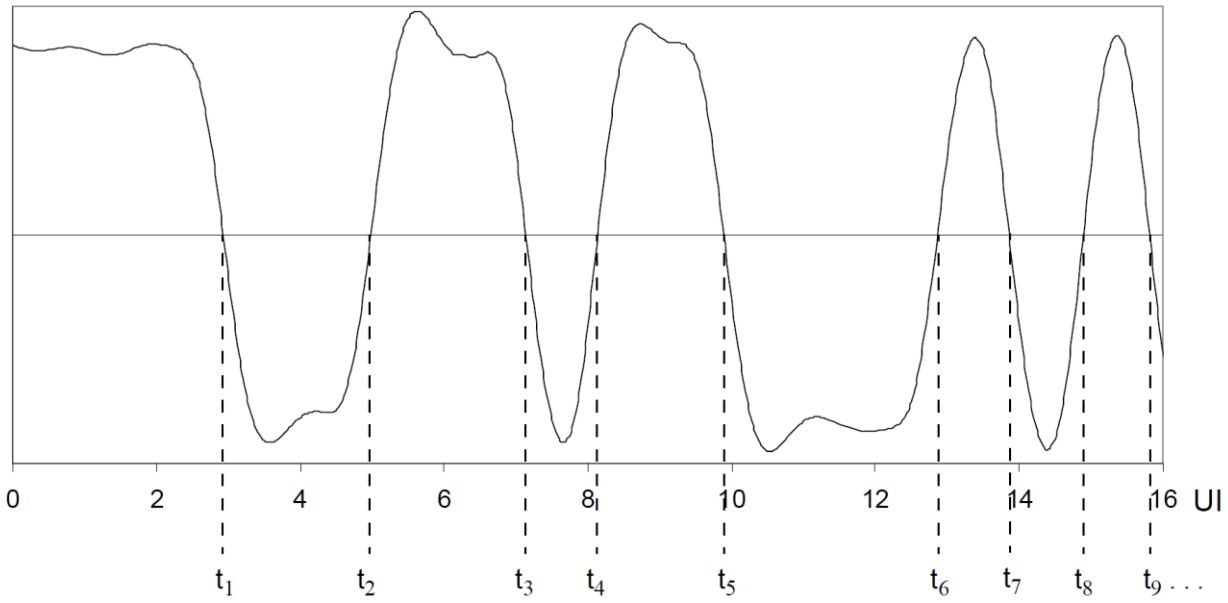


Figure M. 7 *Data Dependent Pulse Width Shrinkage test method.*

The DDPWS is the difference between one symbol period and the minimum of all the differences between pairs of adjacent transitions as follows in **Equation M- 12**.

$$DDPWS = t = \min\left(t_2 - t_1, t_3 - t_2, \ldots\ t_{256+1} - t_{256}\right)$$

Equation M- 12

where *T* is one symbol period.

Note that the difference from the next edge in the repeating sequence, t_{256+1}, is also considered.

M.4.3.5. Signal to noise ratio Q_{sq}

Q_{sq} is a measure of signal to noise ratio. For an electrical signal, it is analogous to Q_{sq} defined for an optical signal, and it relates the low-frequency signal amplitude to the noise in an electrical –3dB bandwidth of 12GHz. It is defined with all co-propagating and counter-propagating crosstalk sources active, using one of patterns 3, 5, or a valid 100GBASE-SR10 signal. The input lanes of the item under test are receiving signals that are asynchronous to those being output.

Q_{sq} may be measured using an oscilloscope as follows:

a. The Voltage Modulation Amplitude (VMA) of the output lane is measured, using a square wave (8 ones, 8 zeros) or PRBS9 (Pattern 4). VMA is the difference

between the 0% level and the 100% level defined in this annex; it is defined by analogy to OMA, but with a 12GHz observation bandwidth.

b. Using the same pattern, the RMS noise over flat regions of the logic one and logic zero portions of the signal, is measured, compensating for noise in the measurement system. If possible, means should be used to prevent noise of frequency less than 1MHz from affecting the result.

c. Q_{sq} is given by **Equation M- 13**.

$$Q_{sq} = \frac{VMA}{(n_1 + n_0)}$$

Equation M- 13

where

n_1 is the RMS noise of logic one

n_0 is the RMS noise of logic zero

M.4.3.6. Eye mask for TP1a and TP4

The eye mask is defined by parameters X1, X2, Y1, and Y2. Unlike the optical eye mask, the vertical dimensions are fixed rather than scaled to the signal. **Figure M. 8** (an example of a hexagonal eye mask such as at TP1a) and **Figure M. 9** (a diamond mask such as at TP4) show the meaning of the parameters X1, X2, Y1 and Y2. The eye is defined as measured using a receiver with an electrical –3dB bandwidth of 12GHz (such as a Bessel-Thomson response). Further requirements are given in Section **2.6.4**.

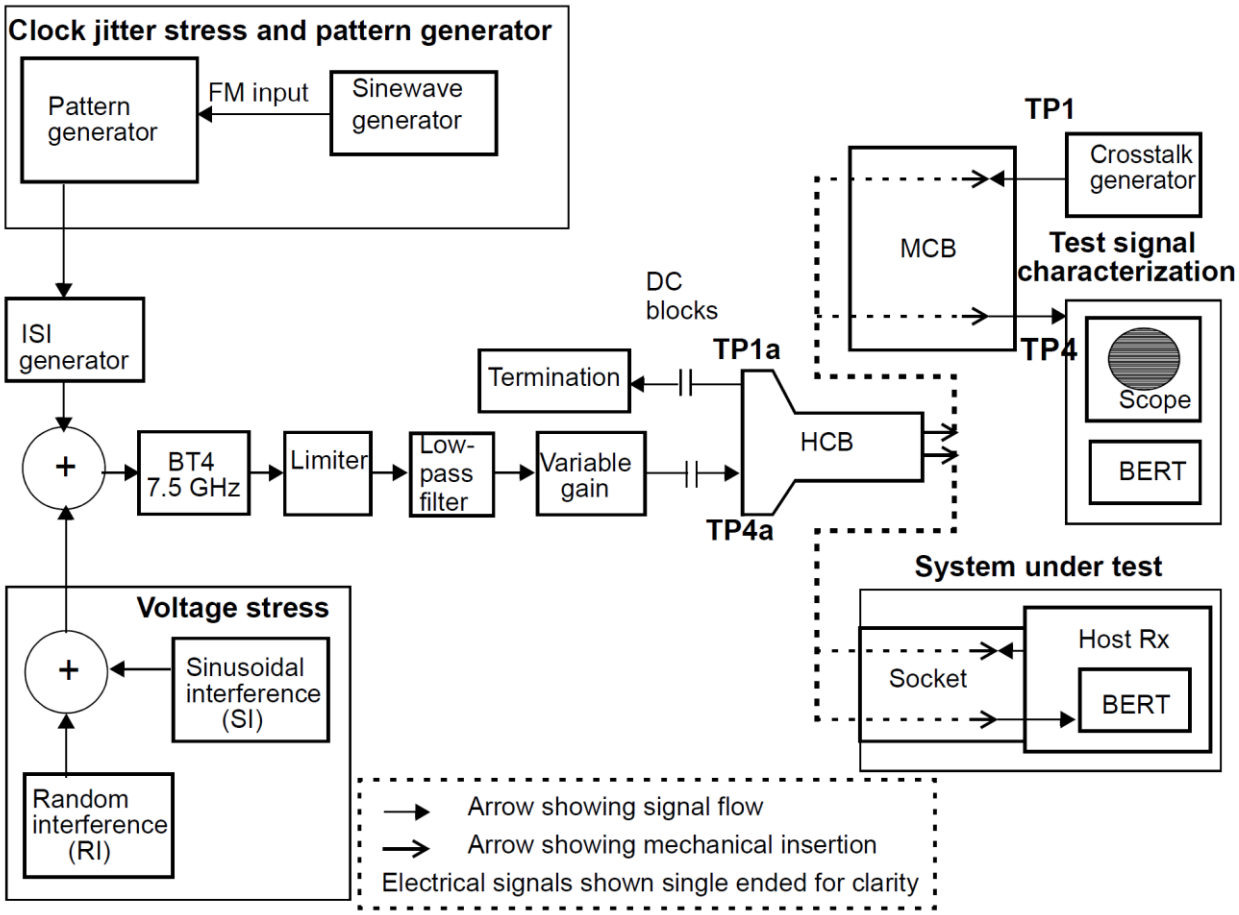


Figure M. 8. *Example jitter tolerance test configuration*

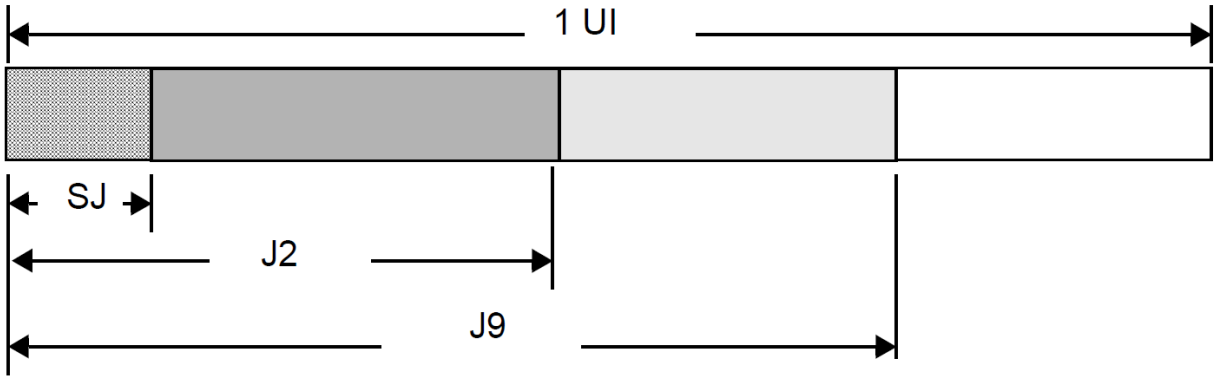


Figure M. 9. *Stressed eye jitter components*

5 References

D'Ambrosia, J., & Nowell, M. (2022). *Proposed Objectives – P802.3dj PAR.* IEEE.

D'Ambrosia, J., & Nowell, M. (2022). *Proposed Objectives –Modified P802.3df PAR.* IEEE.

D'Ambrosia, J., & Nowell, M. (2022). *Splitting IEEE P802.3df, Proposed Draft IEEE P802.3df PAR Modification Responses 400 GbE and 800 GbE Objectives.* IEEE.

D'Ambrosia, J., & Nowell, M. (2022). *Splitting IEEE P802.3df, Proposed Draft IEEE P802.3dj, PAR Responses200 GbE, 400 GbE, 800 GbE and 1.6 TbE Objectives.* IEEE.

D'Ambrosia, J., Law, D., & Nowell, M. (2010). *40 Gigabit Ethernet and 100 Gigabit Ethernet. Technology Overview.* Ethernet Alliance.

D'Ambrosia, J. (2021). *Draft Project Documentation Update.* IEEE 802.3 Beyond 400Gb/s Ethernet Study Group.

D'Ambrosia, J. (2021). *Objectives.* IEEE 802.3 Beyond 400Gb/s Ethernet Study Group.

G, M., G, D., & O, E. (2021). *800G Specification.* Ethernet Technology Consotium.

Gustlin, M., Pepper, J., Szczepanek, A., Trowbridge, S., & Wang, T. (2014). *400GbE PCS Options.* IEEE P802.3bs 400 Gb/s Ethernet Task Force.

He, X., Ren, H., & Wang, X. (2021). *FEC Architecture of B400GbE to Support BER Objective.* HUAWEI TECHNOLOGIES CO., LTD.

IEEE STANDARDS ASSOCATION. (2017). *Amendment 10: Media Access Control Parameters, Physical Layers, and Management Parameters for 200 Gb/s and 400 Gb/s Operation.* IEEE.

IEEE Std 802.3™. (2018). IEEE.

ITU-T. (2004). *G.975.1 - Forward error correction for high bit-rate DWDM submarine systems.* ITU-T.

LU, Y., & Zhuang, Y. (2021). *Considerations on 200G per lane PAM Signaling.* Huawei Technologies.

LU, Y., & ZHUANG, Y. (2021). *Considerations on beyond 100G per lane electrical objectives.* IEEE P802.3 B400G Study Group.

LU, Y., & Zhuang, Y. (2022). *DSP and FEC Considerations for 800GbE and 1.6TbE.* IEEE.

Lyubomirsky, I. (2021). *PAM4 and FEC Architectures for Beyond 400Gb/s Data Center Interconnects.* ISSCC Forum.

Nagarajan, R., Lyubomirsky, I., & Agazzi, O. (2021). *Low Power DSP-Based Transceivers for Data Center.* Journal of Lightwave Technology, VOL. 39, NO. 16.

Nowell, M., & Ofelt, D. (2021). *Thoughts on the Beyond 400 GbE Study Group.* Cisco - Juniper.

Ran, A. (2021). *Addressing possible 800G copper cable objective.* Cisco.

Shrikhande, K. (2021). *Enabling Dense 200GbE and 400GbE.* Innovium.

Theodoras, J. (2021). *PMA lane versus PCS frame FEC and the resulting impact on guaranteeing performance at transceiver module / host PCB interface.* HG Genuine.

Wang, X., & He, X. (2022). *Analysis on FEC1 in Archtecture of 800G/1.6TbE.* IEEE.

Wang, X., & He, X. (2022). *FEC Code and Scheme Observation in 800G/1.6TbE.* IEEE.

Wang, X., He, X., & Ren, H. (2021). *BER objective for Beyond 400GbE.* HUAWEI.

Wang, X., He, X., & Ren, H. (2021). *Technical Feasibility of Logic Layer to Support Rate Objective.* HUAWEI.

Wang, X., Wang, T., & Yang, W. (2014). *Technical Feasibility of Bit multiplexing in 400GbE PMA.* HUAWEI TECHNOLOGIES CO., LTD.