IEEE P802.3bj™/D1.1
Draft Standard for Ethernet

Amendment X: Physical Layer and Management Parameters for 100 Gb/s Operation Over Backplanes and Copper Cables

Sponsor

LAN/MAN Standards Committee
of the
IEEE Computer Society

Approved XX Month 201X

IEEE-SA Standard Board

This draft is an amendment of IEEE Std 802.3-201X. Draft D3.1 and is prepared by the IEEE P802.3bj 100 Gb/s Backplane and Copper Cable Task Force for Task Force review. This draft expires 6 months after the date of publication or when the next version is published, whichever comes first.

IEEE Standards Activities Department
Manager, Standards Intellectual Property
445 Hoes Lane, P.O. Box 1331
Piscataway, NJ 08855-1331, USA

Abstract: [insert abstract]

Keywords: [insert keyword. examples include the port names]

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

PDF: ISBN 973-07381-xxxx-x STDxxxxx
Print: ISBN 973-07381-xxxx-x STDPDxxxxx

IEEE Standards documents are developed within the IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (IEEE-SA) Standards Board. The IEEE develops its standards through a consensus development process, approved by the American National Standards Institute, which brings together volunteers representing varied viewpoints and interests to achieve the final product. Volunteers are not necessarily members of the Institute and serve without compensation. While the IEEE administers the process and establishes rules to promote fairness in the consensus development process, the IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

Use of an IEEE Standard is wholly voluntary. The IEEE disclaims liability for any personal injury, property or other damage, of any nature whatsoever, whether special, indirect, consequential, or compensatory, directly or indirectly resulting from the publication, use of, or reliance upon this, or any other IEEE Standard document.

The IEEE does not warrant or represent the accuracy or content of the material contained herein, and expressly disclaims any express or implied warranty, including any implied warranty of merchantability or fitness for a specific purpose, or that the use of the material contained herein is free from patent infringement. IEEE Standards documents are supplied "**AS IS**."

The existence of an IEEE Standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE Standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard. Every IEEE Standard is subjected to review at least every five years for revision or reaffirmation. When a document is more than five years old and has not been reaffirmed, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE Standard (refer to sections titled "Updating of IEEE documents" and "Errata").

In publishing and making this document available, the IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity. Nor is the IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing this, and any other IEEE Standards document, should rely upon his or her independent judgement in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that his or her views should be considered the personal views of that individual rather than the formal position, explanation, or interpretation of the IEEE.

Comments for revision of IEEE Standards are welcome from any interested party, regardless of membership affiliation with IEEE. Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Comments on standards and requests for interpretations should be submitted to the following address:

Secretary, IEEE-SA Standards Board
445 Hoes Lane
Piscataway, NJ 08854
USA

Introduction

This introduction is not part of IEEE Std 802.3-201X, IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements, Part 3: CSMA/CD Access Method and Physical Layer Specifications.

IEEE Std 802.3™ was first published in 1985. Since the initial publication, many projects have added functionality or provided maintenance updates to the specifications and text included in the standard. Each IEEE 802.3 project/amendment is identified with a suffix (e.g., IEEE Std 802.3ba™-2010). A historical listing of all projects that have added to or modified IEEE Std 802.3 follows as a part of this introductory material. The listing is in chronological order of project initiation and for each project describes: subject, clauses added (if any), approval dates, and committee officers.

The Media Access Control (MAC) protocol specified in IEEE Std 802.3 is Carrier Sense Multiple Access with Collision Detection (CSMA/CD). This MAC protocol was included in the experimental Ethernet developed at Xerox Palo Alto Research Center. While the experimental Ethernet had a 2.94 Mb/s data rate, IEEE Std 802.3-1985 specified operation at 10 Mb/s. Since 1985 new media options, new speeds of operation, and new capabilities have been added to IEEE Std 802.3.

Some of the major additions to IEEE Std 802.3 are identified in the marketplace with their project number. This is most common for projects adding higher speeds of operation or new protocols. For example, IEEE Std 802.3u™ added 100 Mb/s operation (also called Fast Ethernet), IEEE Std 802.3x specified full duplex operation and a flow control protocol, IEEE Std 802.3z added 1000 Mb/s operation (also called Gigabit Ethernet), IEEE Std 802.3ae added 10 Gb/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 40 Gb/s operation (also called 40 Gigabit Ethernet) and 100 Gb/s operation (also called 100 Gigabit Ethernet). These major additions are all now included in and are superseded by IEEE Std 802.3-201X and are not maintained as separate documents.

IEEE Std 802.3-201X

Section One—Includes Clause 1 through Clause 20 and Annex A through Annex H and Annex 4A. Section One includes the specifications for 10 Mb/s operation and the MAC, frame formats and service interfaces used for all speeds of operation.

Section Two—Includes Clause 21 through Clause 33 and Annex 22A through Annex 33E. Section Two includes management attributes for multiple protocols and speed of operation as well as specifications for providing power over twisted pair cabling for multiple operational speeds. It also includes general information on 100 Mb/s operation as well as most of the 100 Mb/s Physical Layer specifications.

Section Three—Includes Clause 34 through Clause 43 and Annex 36A through Annex 43C. Section Three includes general information on 1000 Mb/s operation as well as most of the 1000 Mb/s Physical Layer specifications.

Section Four—Includes Clause 44 through Clause 55 and Annex 44A through Annex 55B. Section Four includes general information on 10 Gb/s operation as well as most of the 10 Gb/s Physical Layer specifications.

Section Five—Includes Clause 56 through Clause 77 and Annex 57A through Annex 76A. Clause 56 through Clause 67 and Clause 75 through Clause 77, as well as associated annexes, specify subscriber access and other Physical Layers and sublayers for operation from 512 kb/s to 10 Gb/s, and defines services and protocol elements that enable the exchange of IEEE Std 802.3 format frames between sta-

tions in a subscriber access network. Clause 68 specifies a 10 Gb/s Physical Layer specification. Clause 69 through Clause 74 and associated annexes specify Ethernet operation over electrical backplanes at speeds of 1000 Mb/s and 10 Gb/s.

Section Six—Includes Clause 78 through Clause 90 and Annex 83A through Annex 86A. Clause 78 specifies Energy-Efficient Ethernet. Clause 79 specifies IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements. Clause 80 through Clause 89 and associated annexes includes general information on 40 Gb/s and 100 Gb/s operation as well the 40 Gb/s and 100 Gb/s Physical Layer specifications. Clause 90 specifies Ethernet support for time synchronization protocols.

IEEE Std 802.3 will continue to evolve. New Ethernet capabilities are anticipated to be added within the next few years as amendments to this standard.

Notice to users

Laws and regulations

Users of these documents should consult all applicable laws and regulations. Compliance with the provisions of this standard does not imply compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Copyrights

Updating of IEEE documents

Users of IEEE standards should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect. In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit the IEEE Standards Association website at http://ieeexplore.ieee.org/xpl/standards.jsp, or contact the IEEE at the address listed previously.

For more information about the IEEE Standards Association or the IEEE standards development process, visit the IEEE-SA website at http://standards.ieee.org.

Errata

Errata, if any, for this and all other standards can be accessed at the following URL: http://standards.ieee.org/findstds/errata/index.html.Users are encouraged to check this URL for errata periodically.

Patents

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE-SA Website <http://standards.ieee.org/about/sasb/patcom/patents.html>. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

Participants

The following individuals were officers of the IEEE 802.3 working group at the beginning of the working group ballot of this revision:

David J. Law, *Working Group Chair*
Wael William Diab, *Working Group Vice-Chair*
Adam Healey, *Working Group Secretary*
Steven B. Carlson, *Working Group Executive Secretary*
Valerie Maguire, *Working Group Treasurer*

John D'Ambrosia, *IEEE P802.3bj Task Force Chair*
Adam Healey, *IEEE P802.3bj Task Force Editor-in-Chief*

The following members of the individual balloting committee voted on this standard. Balloters may have voted for approval, disapproval, or abstention.

When the IEEE–SA Standards Board approved this standard on xx Month 201x, it had the following membership:

FirstName SecondName, *Chair*
FirstName SecondName, *Vice Chair*
FirstName SecondName, *Past Chair*
Judith Gorman, *Secretary*

[to be supplied at publication]

*Member Emeritus

Also included are the following nonvoting IEEE–SA Standards Board liaisons:

[to be supplied at publication]

Michelle D. Turner
IEEE Standards Senior Program Manager, Document Development

Kathryn M. Bennett
IEEE Standards Program Manager, Technical Program Development

List of special symbols

For the benefit of those who have received this document by electronic means, what follows is a list of special symbols and operators. If any of these symbols or operators fail to print out correctly on your machine, the editors apologize, and hope that this table will at least help you to sort out the meaning of the resulting funny-shaped blobs and strokes

Special symbols and operators

Printed character	Meaning	Font
*	Boolean AND	Symbol
+	Boolean OR, arithmetic addition	Symbol
^	Boolean XOR	Times New Roman
!	Boolean NOT	Symbol
×	Multiplication	Symbol
<	Less than	Symbol
≤	Less than or equal to	Symbol
>	Greater than	Symbol
≥	Greater than or equal to	Symbol
=	Equal to	Symbol
≈	Approximately equal to	Symbol
≠	Not equal to	Symbol
⇐	Assignment operator	Symbol
∈	Indicates membership	Symbol
∉	Indicates nonmembership	Symbol
±	Plus or minus (a tolerance)	Symbol
°	Degrees	Symbol
∑	Summation	Symbol
√	Square root	Symbol
—	Big dash (em dash)	Times New Roman
–	Little dash (en dash), subtraction	Times New Roman
|	Vertical bar	Times New Roman
†	Dagger	Times New Roman
‡	Double dagger	Times New Roman
α	Lower case alpha	Symbol
β	Lower case beta	Symbol
γ	Lower case gamma	Symbol
δ	Lower case delta	Symbol
ε	Lower case epsilon	Symbol
λ	Lower case lambda	Symbol
μ	Lower case mu	Times New Roman
Π	Upper case pi	Symbol
Ω	Upper case omega	Symbol

Contents

1. Introduction ... 18

1.4 Definitions ... 18

30. Management ... 19

45. Management Data Input/Output (MDIO) Interface ... 21

45.2.7.12 Backplane Ethernet, BASE-R copper status (Register 7.48) ... 21
45.2.7.13 EEE advertisement (Register 7.60) ... 22
45.2.7.14 EEE link partner ability (Register 7.61) ... 25
45.5 Protocol implementation conformance statement (PICS) proforma for Clause 45, Management Data Input/Output (MDIO) Interface ... 27

69. Introduction to Ethernet operation over electrical backplanes ... 28

69.1.1 Scope ... 28
69.1.2 Objectives ... 28
69.1.3 Relationship of Backplane Ethernet to the ISO OSI reference model ... 28
69.2.1 Reconciliation sublayer and media independent interfaces ... 30
69.2.4 Physical Layer signaling systems ... 31
69.2.6 Low-Power Idle ... 32
69.3 Delay constraints ... 32
69.5 Protocol implementation conformance statement (PICS) proforma ... 32

73. Auto-Negotiation for backplane and copper cable assembly ... 33

73.3 Functional specifications ... 33
73.5.1 DME electrical specifications ... 33
73.6.4 Technology Ability Field ... 33
73.7 Receive function requirements ... 34
73.7.1 DME page reception ... 34
73.7.2 Receive Switch function ... 34
73.7.6 Priority Resolution function ... 34
73.10.7 State diagram variables ... 35
73.11 Protocol implementation conformance statement (PICS) proforma for Clause 73, Auto-Negotiation for backplane and copper cable assembly ... 36

78. Energy-Efficient Ethernet ... 37

78.1 Overview ... 37
78.1.4 EEE Supported PHY types ... 38
78.2 LPI mode timing parameters description ... 38
78.5 Communication link access latency ... 38
78.5.2 100 Gb/s PHY extension using CAUI ... 39
78.6 Protocol implementation conformance statement (PICS) proforma for EEE Data Link Layer Capabilities ... 41

80. Introduction to 40 Gb/s and 100 Gb/s networks ... 42

80.1.2 Objectives ... 42

80.1.3 Relationship of 40 Gigabit and 100 Gigabit Ethernet to the ISO OSI reference model 42
80.1.4 Nomenclature 43
80.1.5 Physical Layer signaling systems 44
80.2.3 Forward Error Correction (FEC) sublayers 46
80.2.4 Physical Medium Attachment (PMA) sublayer 46
80.2.5 Physical Medium Dependent (PMD) sublayer 46
80.2.6 Auto-Negotiation 46
80.3.1 Inter-sublayer service interface 46
80.3.2 Inter-sublayer service interface 46
80.3.3.4 IS_TX_MODE.request 47
80.3.3.5 IS_RX_MODE.indicate 47
80.4 Delay constraints 50
80.5 Skew constraints 50
80.7 Protocol implementation conformance statement (PICS) proforma for Clause 80, Introduction to 40 Gb/s and 100 Gb/s networks 54

81. Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation (XLGMII and CGMII) 55

81.1 Overview 55
81.1.1 Summary of major concepts 55
81.1.7 Mapping of XLGMII/CGMII signals to PLS service primitives 55
81.1.7.3 Mapping of PLS_CARRIER.indication 56
81.3 XLGMII/CGMII functional specifications 56
81.3.1.2 TXC<7:0> (transmit control) 56
81.3.1.5 Transmit direction LPI transition 56
81.3.2.2 RXC<7:0> (receive control) 57
81.3.2.4 Receive direction LPI transition 57
81.3.4 Link fault signaling 58
81.3a LPI Assertion and Detection 59
81.3a.1 LPI messages 59
81.3a.2 Transmit LPI state diagram 60
81.3a.2.1 Variables and counters 60
81.3a.2.2 State Diagram 60
81.3a.3 Considerations for transmit system behavior 60
81.3a.3.1 Considerations for receive system behavior 61
81.4 Protocol implementation conformance statement (PICS) proforma for Clause 81, Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation 62
81.4.2.3 Major capabilities/options 62
81.4.3.6 LPI functions 62

82. Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R 63

82.1.3 Summary of 40GBASE-R and 100GBASE-R sublayers 63
82.1.4 Inter-sublayer interfaces 63
82.1.4.1 PCS service interface (XLGMII/CGMII) 63
82.1.4.2 Physical Medium Attachment (PMA) or Forward Error Correction (FEC) service interface 64
82.1.5 Functional block diagram 64
82.2.3.4 Control codes 64
82.2.3.6 Idle (/I/) 64
82.2.8a Rapid alignment marker insertion 66

82.2.11 Block synchronization 67
82.2.12 PCS lane deskew 67
82.2.18.3 State diagrams 71
82.3.1 PMD MDIO function mapping 72
82.6 Auto-Negotiation 72
82.7 Protocol implementation conformance statement (PICS) proforma for Clause 82, Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R 81
82.7.3 Major Capabilities/Options 81
82.7.6.5 Auto-Negotiation for Backplane Ethernet functions 81
82.7.6.6 LPI functions 82

83. Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R 83

83.1 Overview 83
83.1.1 Scope 83
83.3 PMA service interface 83
83.6 PMA MDIO function mapping 84
83.7 Protocol implementation conformance statement (PICS) proforma for Clause 83, Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R 85
83.7.3 Major Capabilities/Options 85

84. Physical Medium Dependent sublayer and baseband medium, type 40GBASE-KR4 86

85. Physical Medium Dependent sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10 87

85.1 Overview 87
85.2 Physical Medium Dependent (PMD) service interface 87
85.7.2 PMD Transmit function 88
85.7.4 Global PMD signal detect function 88
85.7.6 Global PMD transmit disable function 88
85.8.3 Transmitter characteristics 89
85.13 Protocol implementation conformance statement (PICS) proforma for Clause 85, Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10 90
85.13.3 Major capabilities/options 90

91. Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs 91

91.1 Overview 91
91.1.1 Scope 91
91.1.2 Position of RS-FEC in the 100GBASE-R sublayers 91
91.2 FEC service interface 91
91.3 PMA compatibility 92
91.4 Delay constraints 92
91.5 Functions within the RS-FEC sublayer 93
91.5.1 Functional block diagram 93
91.5.2 Transmit function 93
91.5.2.1 Lane block synchronization 93
91.5.2.2 Alignment lock and deskew 93
91.5.2.3 Lane reorder 93
91.5.2.4 Alignment marker removal 93
91.5.2.5 64B/66B to 256B/257B transcoder 94
91.5.2.6 Alignment mapping and insertion 95

91.5.2.7 Reed-Solomon encoder ... 97
91.5.2.8 Symbol distribution ... 99
91.5.2.9 Transmit bit ordering ... 99
91.5.3 Receive function ... 99
91.5.3.1 Alignment lock and deskew ... 99
91.5.3.2 Lane reorder ... 99
91.5.3.3 Reed-Solomon decoder ... 101
91.5.3.4 Alignment marker removal ... 101
91.5.3.5 256B/257B to 64B/66B transcoder ... 101
91.5.3.6 Block distribution ... 102
91.5.3.7 Alignment marker mapping and insertion ... 102
91.5.3.8 Receive bit ordering ... 102
91.5.4 Detailed functions and state diagrams ... 104
91.5.4.1 State diagram conventions ... 104
91.5.4.2 State variables ... 104
91.5.4.3 State diagrams ... 106
91.6 FEC management ... 108
91.7 Protocol implementation conformance statement (PICS) proforma for Clause 91, Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs ... 109
91.7.1 Introduction ... 109
91.7.2 Identification ... 109
91.7.2.1 Implementation identification ... 109
91.7.2.2 Protocol summary ... 109

92. Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-CR4 ... 111

92.1 Overview ... 111
92.2 Physical Medium Dependent (PMD) service interface ... 112
92.3 PCS requirements for Auto-Negotiation (AN) service interface ... 113
92.4 Delay constraints ... 113
92.5 Skew constraints ... 114
92.6 PMD MDIO function mapping ... 114
92.7 PMD functional specifications ... 114
92.7.1 Link block diagram ... 114
92.7.2 PMD Transmit function ... 117
92.7.3 PMD Receive function ... 117
92.7.4 Global PMD signal detect function ... 117
92.7.5 PMD lane-by-lane signal detect function ... 117
92.7.6 Global PMD transmit disable function ... 117
92.7.7 PMD lane-by-lane transmit disable function ... 118
92.7.8 Loopback mode ... 118
92.7.9 PMD_fault function ... 118
92.7.10 PMD transmit fault function ... 118
92.7.11 PMD receive fault function ... 118
92.7.12 PMD control function ... 119
92.8 100GBASE-CR4 electrical characteristics ... 119
92.8.1 Signal levels ... 119
92.8.2 Signal paths ... 119
92.8.3 Transmitter characteristics ... 120
92.8.3.1 Transmitter differential output return loss ... 120
92.8.3.2 Transmitter noise parameter measurements ... 121
92.8.3.3 Transmitter output waveform ... 121
92.8.3.4 Insertion loss TP0 to TP2 or TP3 to TP5 ... 126
92.8.3.5 Test fixture ... 127

92.8.3.6 Test fixture impedance ... 128
92.8.3.7 Test fixture insertion loss ... 128
92.8.3.8 Transmitter output jitter ... 128
92.8.3.9 Signaling rate range ... 129
92.8.4 Receiver characteristics at TP3 summary ... 130
92.8.4.1 Receiver differential input return loss ... 130
92.8.4.2 Receiver interference tolerance test ... 131
92.8.4.3 Bit error ratio ... 133
92.8.4.4 Signaling rate range ... 133
92.8.4.5 AC coupling ... 133
92.9 Channel characteristics ... 133
92.10 Cable assembly characteristics ... 133
92.10.1 Characteristic impedance and reference impedance ... 134
92.10.2 Cable assembly insertion loss ... 134
92.10.3 Cable assembly insertion loss deviation (ILD) ... 136
92.10.4 Cable assembly return loss ... 136
92.10.5 Cable assembly multiple disturber near-end crosstalk (MDNEXT) loss ... 137
92.10.6 Cable assembly multiple disturber far-end crosstalk (MDFEXT) loss ... 138
92.10.7 Cable assembly integrated crosstalk noise (ICN) ... 138
92.10.8 Cable assembly test fixture ... 140
92.10.9 Mated test fixtures ... 141
92.10.9.1 Mated test fixtures insertion loss ... 141
92.10.9.2 Mated test fixtures return loss ... 142
92.10.9.3 Mated test fixtures common-mode conversion loss ... 143
92.10.9.4 Mated test fixtures integrated crosstalk noise ... 144
92.11 MDI specification ... 145
92.12 Environmental specifications ... 146
92.13 Protocol implementation conformance statement (PICS) proforma for Clause 92, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-CR4 ... 147
92.13.1 Introduction ... 147
92.13.2 Identification ... 147
92.13.2.1 Implementation identification ... 147
92.13.2.2 Protocol summary ... 147

93. Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4 ... 149

93.1 Overview ... 149
93.2 Physical Medium Dependent (PMD) service interface ... 150
93.3 PCS requirements for Auto-Negotiation (AN) service interface ... 151
93.4 Delay constraints ... 151
93.5 Skew constraints ... 152
93.6 PMD MDIO function mapping ... 152
93.7 PMD functional specifications ... 154
93.7.1 Link block diagram ... 154
93.7.2 PMD Transmit function ... 154
93.7.3 PMD Receive function ... 154
93.7.4 Global PMD signal detect function ... 154
93.7.5 PMD lane-by-lane signal detect function ... 155
93.7.6 Global PMD transmit disable function ... 155
93.7.7 PMD lane-by-lane transmit disable function ... 155
93.7.8 Loopback mode ... 155
93.7.9 PMD_fault function ... 156
93.7.10 PMD transmit fault function ... 156
93.7.11 PMD receive fault function ... 156

93.7.12 PMD control function ... 156
93.8 100GBASE-KR4 electrical characteristics ... 156
93.8.1 Transmitter characteristics ... 156
93.8.1.1 Transmitter test fixture ... 156
93.8.1.2 Signaling rate and range ... 158
93.8.1.3 Signal levels ... 158
93.8.1.4 Transmitter output return loss ... 158
93.8.1.5 Transition time ... 158
93.8.1.6 Transmitter output waveform ... 159
93.8.1.7 Transmitter far-end output noise ... 161
93.8.1.8 Transmitter output jitter ... 161
93.8.2 Receiver characteristics ... 162
93.8.2.1 Receiver test fixture ... 162
93.8.2.2 Receiver input return loss ... 162
93.8.2.3 Receiver interference tolerance ... 163
93.8.3 AC-coupling ... 163
93.9 Channel characteristics ... 164
93.9.1 Channel operating margin ... 164
93.9.2 Insertion loss ... 164
93.9.3 Channel Integrated Crosstalk Noise (ICN) ... 166
93.10 Environmental specifications ... 166
93.10.1 General safety ... 166
93.10.2 Network safety ... 166
93.10.3 Installation and maintenance guidelines ... 166
93.10.4 Electromagnetic compatibility ... 166
93.10.5 Temperature and humidity ... 166
93.11 Protocol implementation conformance statement (PICS) proforma for Clause 93, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4 ... 167
93.11.1 Introduction ... 167
93.11.2 Identification ... 167
93.11.2.1 Implementation identification ... 167
93.11.2.2 Protocol summary ... 167
93.11.3 Major capabilities/options ... 168
93.11.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4 ... 168

94. Physical Medium Attachment (PMA) sublayer, Physical Medium Dependent (PMD) sublayer, and baseband medium, type 100GBASE-KP4 ... 169

94.1 Overview ... 169
94.2 Physical Medium Attachment (PMA) Sublayer ... 171
94.2.1 PMA Service Interface ... 171
94.2.1.1 PMA:IS_UNITDATA_i.request ... 171
94.2.1.2 PMA:IS_UNITDATA_i.indication ... 172
94.2.1.3 PMA:IS_SIGNAL.indication ... 172
94.2.1.4 Effect of receipt ... 173
94.2.2 PMA Transmit Functional Specifications ... 173
94.2.2.1 FEC Interface ... 173
94.2.2.2 Overhead Frame ... 174
94.2.2.3 Overhead ... 174
94.2.2.4 Termination Blocks ... 174
94.2.2.5 Gray Mapping ... 175
94.2.2.6 Precoding ... 175
94.2.2.7 PAM4 encoding ... 175

94.2.2.8 PMD Interface ... 176
94.2.3 PMA Transmit EEE Operation ... 176
94.2.4 PMA Receive Functional Specifications ... 176
94.2.5 PMA Receive EEE Operation ... 177
94.2.6 Skew constraints ... 177
94.2.7 Delay constraints ... 177
94.2.8 Link status ... 178
94.2.9 PMA local loopback mode ... 178
94.2.10 PMA remote loopback mode (optional) ... 178
94.2.11 PMA test patterns ... 179
94.2.12 PMA MDIO function mapping ... 179
94.3 Physical Medium Dependent (PMD) Sublayer ... 179
94.3.1 Physical Medium Dependent (PMD) service interface ... 179
94.3.1.1 PMD:IS_UNITDATA_i.request ... 180
94.3.1.2 PMD:IS_UNITDATA_i.indication ... 180
94.3.1.3 PMD:IS_SIGNAL.indication ... 181
94.3.2 PCS requirements for Auto-Negotiation (AN) service interface ... 181
94.3.3 Delay constraints ... 182
94.3.4 Skew constraints ... 182
94.3.5 PMD MDIO function mapping ... 182
94.3.6 PMD functional specifications ... 184
94.3.6.1 Link block diagram ... 184
94.3.6.2 PMD Transmit function ... 184
94.3.6.3 PMD Receive function ... 184
94.3.6.4 Global PMD signal detect function ... 184
94.3.6.5 PMD lane-by-lane signal detect function ... 185
94.3.6.6 Global PMD transmit disable function ... 185
94.3.6.7 PMD lane-by-lane transmit disable function ... 185
94.3.6.8 Loopback mode ... 185
94.3.7 PMD_fault function ... 186
94.3.8 PMD transmit fault function ... 186
94.3.9 PMD receive fault function ... 186
94.3.10 PMD control function ... 186
94.3.11 PMD Transmitter electrical characteristics ... 187
94.3.11.1 Test Fixture ... 188
94.3.11.2 Signaling rate and range ... 188
94.3.11.3 Signal levels ... 188
94.3.11.4 Transmitter output return loss ... 189
94.3.11.5 Transition time ... 189
94.3.11.6 Transmitter linearity ... 189
94.3.11.7 Transmitter output waveform ... 191
94.3.11.8 Transmitter far-end output noise ... 193
94.3.11.9 Transmitter output jitter ... 193
94.3.12 PMD Receiver electrical characteristics ... 194
94.3.12.1 Test Fixture ... 194
94.3.12.2 Receiver input return loss ... 195
94.3.12.3 Receiver interference tolerance ... 195
94.3.13 AC-coupling ... 196
94.4 Channel characteristics ... 196
94.4.1 Channel operating margin ... 196
94.4.2 Channel insertion loss ... 196
94.4.3 Channel Integrated Crosstalk Noise (ICN) ... 198
94.5 Environmental specifications ... 198
94.5.1 General safety ... 198

94.5.2 Network safety ... 198
94.5.3 Installation and maintenance guidelines ... 198
94.5.4 Electromagnetic compatibility ... 199
94.5.5 Temperature and humidity ... 199
94.6 Protocol implementation conformance statement (PICS) proforma for Clause 94, Physical Medium Attachment (PMA) and Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KP4 ... 200
94.6.1 Introduction ... 200
94.6.2 Identification ... 200
94.6.2.1 Implementation identification ... 200
94.6.2.2 Protocol summary ... 200
94.6.3 Major capabilities/options ... 201
94.6.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KP4 ... 201

Annex 83A (normative) 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI) ... 202

83A.3.2a EEE operation ... 202
83A.3.3.6 Global transmit disable function ... 203
83A.3.4.7 Global energy detect function ... 203
83A.4 Protocol implementation conformance statement (PICS) proforma for Annex 83A, 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI) ... 204

Annex 83C (informative) PMA sublayer partitioning examples ... 205

83C.1 Partitioning examples with FEC ... 205
83C.1.1 FEC implemented with PCS ... 205
83C.1.2 FEC implemented with PMD ... 206
83C.1a Partitioning examples with RS FEC ... 207
83C.1a.1 Single PMA sublayer with RS FEC ... 207
83C.1a.2 Single CAUI with RS FEC ... 207

Annex 92A (informative) 100GBASE-CR4 TP0 and TP5 test point parameters ... 208

92A.1 Overview ... 208
92A.2 Transmitter characteristics at TP0 ... 208
92A.3 Receiver characteristics at TP5 ... 208
92A.4 Transmitter and receiver differential printed circuit board trace loss ... 208
92A.5 Channel insertion loss ... 209
92A.6 Channel return loss ... 210
92A.7 Channel insertion loss deviation (ILD) ... 211
92A.8 Channel integrated crosstalk noise (ICN) ... 211

Annex 93A (normative) Characteristics of electrical backplanes ... 213

93A.1 Channel operating margin ... 213
93A.1.1 Measurement of the channel ... 213
93A.1.2 Path terminations ... 215
93A.1.3 Filters ... 215
93A.1.4 Single bit response ... 216
93A.1.5 Signal amplitude ... 216
93A.1.6 Interference and noise amplitude ... 217
93A.1.6.1 Interference amplitude distribution ... 217

93A.1.6.2 Noise amplitude distribution 218
93A.1.6.3 Combination of interference and noise distributions 219

Annex 93B (informative) Electrical backplane reference model 220

1. Introduction

1.4 Definitions

Insert the following definition after 1.4.51:

1.4.51a 100GBASE-CR4: IEEE 802.3 Physical Layer specification for 100 Gb/s using 100GBASE-R encoding and Clause 91 RS-FEC over four lanes of shielded balanced copper cabling, with reach up to at least 5 m. (See IEEE Std 802.3, Clause 92.)

Insert the following definitions after 1.4.53:

1.4.53a 100GBASE-KP4: IEEE 802.3 Physical Layer specification for 100 Gb/s using 100GBASE-R encoding, Clause 91 RS-FEC, and 4-level pulse amplitude modulation over four lanes of an electrical backplane, with a total insertion loss up to 33 dB at 7.0 GHz. (See IEEE Std 802.3, Clause 94.)

1.4.53b 100GBASE-KR4: IEEE 802.3 Physical Layer specification for 100 Gb/s using 100GBASE-R encoding, Clause 91 RS-FEC, and 2-level pulse amplitude modulation over four lanes of an electrical backplane, with a total insertion loss up to 35 dB at 12.9 GHz. (See IEEE Std 802.3, Clause 93.)

30. Management

Editor's note (to be removed prior to final publication):

Additional amendments will be included as the content of the draft stabilizes.

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

30.6.1.1.5 aAutoNegLocalTechnologyAbility

Insert 100GBASE-CR4, 100GBASE-KR4, and 100GBASE-KP4 after 100GBASE-CR10 as shown:

APPROPRIATE SYNTAX:
A SEQUENCE that meets the requirements of the description below:

global	Reserved for future use
other	See 30.2.5
unknown	Initializing, true state or type not yet known
10BASE-T	10BASE-T half duplex as defined in Clause 14
10BASE-TFD	Full duplex 10BASE-T as defined in Clause 14 and Clause 31
100BASE-T4	100BASE-T4 half duplex as defined in Clause 23
100BASE-TX	100BASE-TX half duplex as defined in Clause 25
100BASE-TXFD	Full duplex 100BASE-TX as defined in Clause 25 and Clause 31
FDX PAUSE	PAUSE operation for full duplex links as defined in Annex 31B
FDX APAUSE	Asymmetric PAUSE operation for full duplex links as defined in Clause 37, Annex 28B, and Annex 31B
FDX SPAUSE	Symmetric PAUSE operation for full duplex links as defined in Clause 37, Annex 28B, and Annex 31B
FDX BPAUSE	Asymmetric and Symmetric PAUSE operation for full duplex links as defined in Clause 37, Annex 28B, and Annex 31B
100BASE-T2	100BASE-T2 half duplex as defined in Clause 32
100BASE-T2FD	Full duplex 100BASE-T2 as defined in Clause 31 and Clause 32
1000BASE-X	1000BASE-X half duplex as specified in Clause 36
1000BASE-XFD	Full duplex 1000BASE-X as specified in Clause 31 and Clause 36
1000BASE-T	1000BASE-T half duplex PHY as specified in Clause 40
1000BASE-TFD	Full duplex 1000BASE-T PHY as specified in Clause 31 and as specified in Clause 40
Rem Fault1	Remote fault bit 1 (RF1) as specified in Clause 37
Rem Fault2	Remote fault bit 2 (RF2) as specified in Clause 37
10GBASE-T	10GBASE-T PHY as specified in Clause 55
1000BASE-KXFD	Full duplex 1000BASE-KX as specified in Clause 70
10GBASE-KX4FD	Full duplex 10GBASE-KX4 as specified in Clause 71

10GBASE-KRFD	Full duplex 10GBASE-KR as specified in Clause 72
40GBASE-KR4	40GBASE-KR4 as specified in Clause 84
40GBASE-CR4	40GBASE-CR4 as specified in Clause 85
100GBASE-CR10	100GBASE-CR10 as specified in Clause 85
100GBASE-CR4	100GBASE-CR4 as specified in Clause 92
100GBASE-KR4	100GBASE-KR4 as specified in Clause 93
100GBASE-KP4	100GBASE-KP4 as specified in Clause 94
Rem Fault	Remote fault bit (RF) as specified in Clause 73
FEC Capable	FEC ability as specified in Clause 73 (see 73.7) and Clause 74
FEC Requested	FEC requested as specified in Clause 73 (see 73.7) and Clause 74
isoethernet	IEEE Std 802.9 ISLAN-16T

45. Management Data Input/Output (MDIO) Interface

Editor's note (to be removed prior to final publication):

Additional registers and bits will be included as the content of the draft stabilizes.

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

45.2.7.12 Backplane Ethernet, BASE-R copper status (Register 7.48)

Change Table 45–189 as shown:

Table 45–189—Backplane Ethernet, BASE-R copper status register (Register 7.48) bit definitions

Bit(s)	Name	Description	RO[a]
7.48.15:~~9~~12	Reserved	Ignore on read	RO
7.48.11	100GBASE-CR4	1 = PMA/PMD is negotiated to perform 100GBASE-CR4 0 = PMA/PMD is not negotiated to perform 100GBASE-CR4	RO
7.48.10	100GBASE-KP4	1 = PMA/PMD is negotiated to perform 100GBASE-KP4 0 = PMA/PMD is not negotiated to perform 100GBASE-KP4	RO

Table 45–189—Backplane Ethernet, BASE-R copper status register (Register 7.48) bit definitions

7.48.9	100GBASE-KR4	1 = PMA/PMD is negotiated to perform 100GBASE-KR4 0 = PMA/PMD is not negotiated to perform 100GBASE-KR4	RO
7.48.8	100GBASE-CR10	1 = PMA/PMD is negotiated to perform 100GBASE-CR10 0 = PMA/PMD is not negotiated to perform 100GBASE-CR10	RO
7.48.7	Reserved	Ignore on read	RO
7.48.6	40GBASE-CR4	1 = PMA/PMD is negotiated to perform 40GBASE-CR4 0 = PMA/PMD is not negotiated to perform 40GBASE-CR4	RO
7.48.5	40GBASE-KR4	1 = PMA/PMD is negotiated to perform 40GBASE-KR4 0 = PMA/PMD is not negotiated to perform 40GBASE-KR4	RO
7.48.4	BASE-R FEC negotiated	1 = PMA/PMD is negotiated to perform BASE-R FEC 0 = PMA/PMD is not negotiated to perform BASE-R FEC	RO
7.48.3	10GBASE-KR	1 = PMA/PMD is negotiated to perform 10GBASE-KR 0 = PMA/PMD is not negotiated to perform 10GBASE-KR	RO
7.48.2	10GBASE-KX4	1 = PMA/PMD is negotiated to perform 10GBASE-KX4 or CX4 0 = PMA/PMD is not negotiated to perform 10GBASE-KX4/CX4	RO
7.48.1	1000BASE-KX	1 = PMA/PMD is negotiated to perform 1000BASE-KX 0 = PMA/PMD is not negotiated to perform 1000BASE-KX	RO
7.48.0	BP AN ability	If a Backplane, BASE-R copper PHY type is implemented, this bit is set to 1	RO

[a]RO = Read only

Change 45.2.7.12.2 as shown:

45.2.7.12.2 Negotiated Port Type (7.48.1, 7.48.2, 7.48.3, 7.48.5, 7.48.6, 7.48.8, 7.48.9, 7.48.10, 7.48.11)

When the AN process has been completed as indicated by the AN complete bit, these bits (1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 40GBASE-KR4, 40GBASE-CR4, 100GBASE-CR10, 100GBASE-KR4, 100GBASE-KP4, 100GBASE-CR4) indicate the negotiated port type. Only one of these bits is set depending on the priority resolution function. System developers need to distinguish between parallel detection of 10GBASE-KX4 and 10GBASE-CX4 based on the MDI and media type present.

45.2.7.13 EEE advertisement (Register 7.60)

Change Table 45–190 as shown:

Table 45–190—EEE advertisement register (Register 7.60) bit definitions

Bit(s)	Name	Description	Clause reference; Next Page bit number	R/W[a]
7.60.15:~~7~~ 14	Reserved	Ignore on read		RO
7.60.13	100GBASE-CR4 EEE	1 = Advertise that the 100GBASE-CR4 has EEE capability 0 = Do not advertise that the 100GBASE-CR4 has EEE capability	TBD	R/W
7.60.12	100GBASE-KP4 EEE	1 = Advertise that the 100GBASE-KP4 has EEE capability 0 = Do not advertise that the 100GBASE-KP4 has EEE capability	TBD	R/W
7.60.11	100GBASE-KR4 EEE	1 = Advertise that the 100GBASE-KR4 has EEE capability 0 = Do not advertise that the 100GBASE-KR4 has EEE capability	TBD	R/W
7.60.10	100GBASE-CR10 EEE	1 = Advertise that the 100GBASE-CR10 has EEE capability 0 = Do not advertise that the 100GBASE-CR10 has EEE capability	TBD	R/W

Table 45–190—EEE advertisement register (Register 7.60) bit definitions

7.60.9	Reserved	Ignore on read		RO
7.60.8	40GBASE-CR4 EEE	1 = Advertise that the 40GBASE-CR4 has EEE capability 0 = Do not advertise that the 40GBASE-CR4 has EEE capability	TBD	R/W
7.60.7	40GBASE-KR4 EEE	1 = Advertise that the 40GBASE-KR4 has EEE capability 0 = Do not advertise that the 40GBASE-KR4 has EEE capability	TBD	R/W
7.60.6	10GBASE-KR EEE	1 = Advertise that the 10GBASE-KR has EEE capability 0 = Do not advertise that the 10GBASE-KR has EEE capability	73.7.7.1; U6	R/W
7.60.5	10GBASE-KX4 EEE	1 = Advertise that the 10GBASE-KX4 has EEE capability 0 = Do not advertise that the 10GBASE-KX4 has EEE capability	73.7.7.1; U5	R/W
7.60.4	1000BASE-KX EEE	1 = Advertise that the 1000BASE-KX has EEE capability 0 = Do not advertise that the 1000BASE-KX has EEE capability	73.7.7.1; U4	R/W
7.60.3	10GBASE-T EEE	1 = Advertise that the 10GBASE-T has EEE capability 0 = Do not advertise that the 10GBASE-T has EEE capability	28.2.3.4.1; U3 / 55.6.1; U24	R/W
7.60.2	1000BASE-T EEE	1 = Advertise that the 1000BASE-T has EEE capability 0 = Do not advertise that the 1000BASE-T has EEE capability	28.2.3.4.1; U2 / 55.6.1; U23	R/W
7.60.1	100BASE-TX EEE	1 = Advertise that the 100BASE-TX has EEE capability 0 = Do not advertise that the 100BASE-TX has EEE capability	28.2.3.4.1; U1 / 55.6.1; U22	R/W
7.60.0	Reserved	Ignore on read		RO

[a]R/W = Read/Write, RO = Read only

Insert 45.2.7.13.1a through 45.2.7.13.1f before 45.2.7.13.1 as follows:

45.2.7.13.1a 100GBASE-CR4 EEE supported (7.60.13)

Support for EEE operation for 100GBASE-CR4, as defined in 92.1, shall be advertized if this bit is set to one.

45.2.7.13.1b 100GBASE-KP4 EEE supported (7.60.12)

Support for EEE operation for 100GBASE-KP4, as defined in 94.1, shall be advertized if this bit is set to one.

45.2.7.13.1c 100GBASE-KR4 EEE supported (7.60.11)

Support for EEE operation for 100GBASE-KR4, as defined in 93.1, shall be advertized if this bit is set to one.

45.2.7.13.1d 100GBASE-CR10 EEE supported (7.60.10)

Support for EEE operation for 100GBASE-CR10, as defined in 85.1, shall be advertized if this bit is set to one.

45.2.7.13.1e 40GBASE-CR4 EEE supported (7.60.8)

Support for EEE operation for 40GBASE-CR4, as defined in 85.1, shall be advertized if this bit is set to one.

45.2.7.13.1f 40GBASE-KR4 EEE supported (7.60.7)

Support for EEE operation for 40GBASE-KR4, as defined in 84.1, shall be advertized if this bit is set to one.

45.2.7.14 EEE link partner ability (Register 7.61)

Change Table 45–191 as shown:

Table 45–191—EEE link partner ability (Register 7.61) bit definitions

Bit(s)	Name	Description	Clause reference; Next Page bit number	R/W[a]
7.61.15:~~7~~ 14	Reserved	Ignore on read		RO
7.61.13	100GBASE-CR4 EEE	1 = Link partner is advertising EEE capability for 10GBASE-CR4 0 = Link partner is not advertising EEE capability for 10GBASE-CR4	TBD	RO
7.61.12	100GBASE-KP4 EEE	1 = Link partner is advertising EEE capability for 10GBASE-KP4 0 = Link partner is not advertising EEE capability for 10GBASE-KP4	TBD	RO
7.61.11	100GBASE-KR4 EEE	1 = Link partner is advertising EEE capability for 10GBASE-KR4 0 = Link partner is not advertising EEE capability for 10GBASE-KR4	TBD	RO
7.61.10	100GBASE-CR10 EEE	1 = Link partner is advertising EEE capability for 10GBASE-CR10 0 = Link partner is not advertising EEE capability for 10GBASE-CR10	TBD	RO
7.61.9	Reserved	Ignore on read		RO
7.61.8	40GBASE-CR4 EEE	1 = Link partner is advertising EEE capability for 40GBASE-CR4 0 = Link partner is not advertising EEE capability for 40GBASE-CR4	TBD	RO

Table 45–191—EEE link partner ability (Register 7.61) bit definitions *(continued)*

7.61.7	40GBASE-KR4 EEE	1 = Link partner is advertising EEE capability for 40GBASE-KR4 0 = Link partner is not advertising EEE capability for 40GBASE-KR4	TBD	RO
7.61.6	10GBASE-KR EEE	1 = Link partner is advertising EEE capability for 10GBASE-KR 0 = Link partner is not advertising EEE capability for 10GBASE-KR	73.7.7.1; U6	RO
7.61.5	10GBASE-KX4 EEE	1 = Link partner is advertising EEE capability for 10GBASE-KX4 0 = Link partner is not advertising EEE capability for 10GBASE-KX4	73.7.7.1; U5	RO
7.61.4	1000BASE-KX EEE	1 = Link partner is advertising EEE capability for 1000BASE-KX 0 = Link partner is not advertising EEE capability for 1000BASE-KX	73.7.7.1; U4	RO
7.61.3	10GBASE-T EEE	1 = Link partner is advertising EEE capability for 10GBASE-T 0 = Link partner is not advertising EEE capability for 10GBASE-T	28.2.3.4.1; U3 / 55.6.1; U24	RO
7.61.2	1000BASE-T EEE	1 = Link partner is advertising EEE capability for 1000BASE-T 0 = Link partner is not advertising EEE capability for 1000BASE-T	28.2.3.4.1; U2 / 55.6.1; U23	RO
7.61.1	100BASE-TX EEE	1 = Link partner is advertising EEE capability for 100BASE-TX 0 = Link partner is not advertising EEE capability for 100BASE-TX	28.2.3.4.1; U1 / 55.6.1; U22	RO
7.61.0	Reserved	Ignore on read		RO

[a]R/W = Read/Write, RO = Read only

45.5 Protocol implementation conformance statement (PICS) proforma for Clause 45, Management Data Input/Output (MDIO) Interface[1]

Editor's note (to be removed prior to final publication):

The PICS proforma will be updated when the content of this clause stabilizes.

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

69. Introduction to Ethernet operation over electrical backplanes

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

69.1.1 Scope

Change the second paragraph of 69.1.1 as shown:

Backplane Ethernet supports the IEEE 802.3 MAC operating at 1000 Mb/s, 10 Gb/s, ~~or~~ 40 Gb/s, or 100 Gb/s. For 1000 Mb/s operation, the family of 1000BASE-X Physical Layer signaling systems is extended to include 1000BASE-KX. For 10 Gb/s operation, two Physical Layer signaling systems are defined. For operation over four logical lanes, the 10GBASE-X family is extended to include 10GBASE-KX4. For serial operation, the 10GBASE-R family is extended to include 10GBASE-KR. For 40 Gb/s operation, there is 40GBASE-KR4 that operates over four lanes. For 100 Gb/s operation, the 100GBASE-R family is extended to include 100GBASE-KR4 and 100GBASE-KP4.

69.1.2 Objectives

Delete 69.1.2.

Editor's note (to be removed prior to final publication):

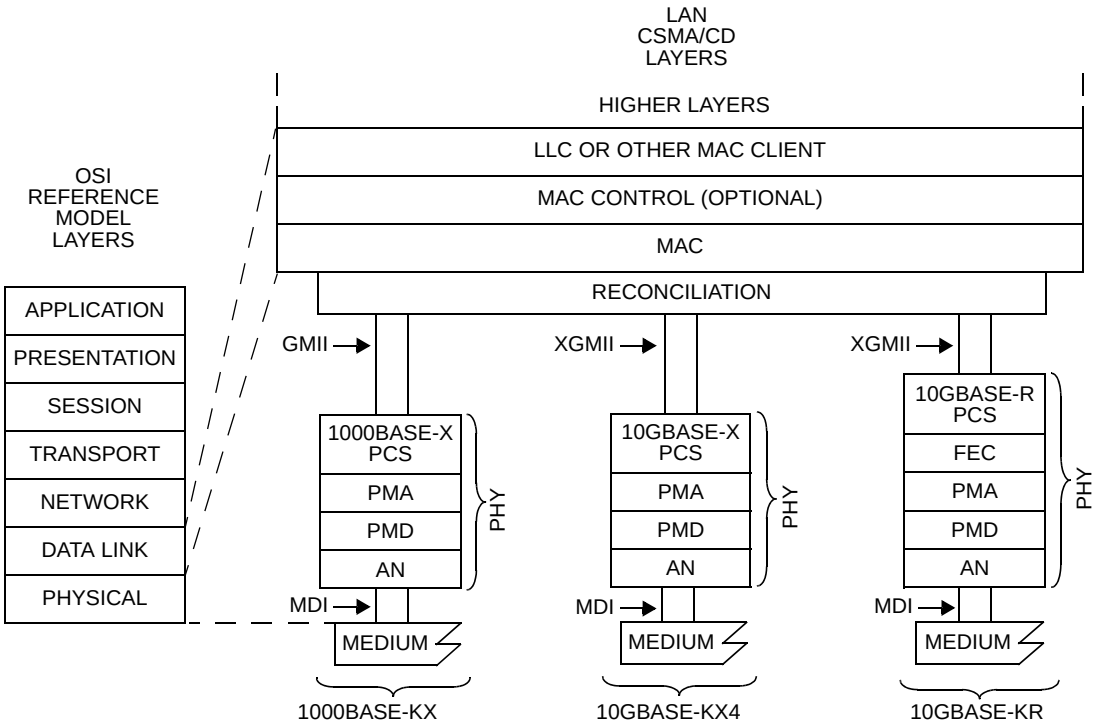
There has been discussion related to the merits of amending lists of objectives such as this as new Physical Layers are defined. This was discussed in the context of comment #20 during the IEEE P802.3bj D1.0 1st Task Force review. The conclusion of that discussion was to delete the objectives subclause of Clause 80. To be consistent, the comparable subclause from Clause 69 is also deleted.

69.1.3 Relationship of Backplane Ethernet to the ISO OSI reference model

Change the first paragraph as shown:

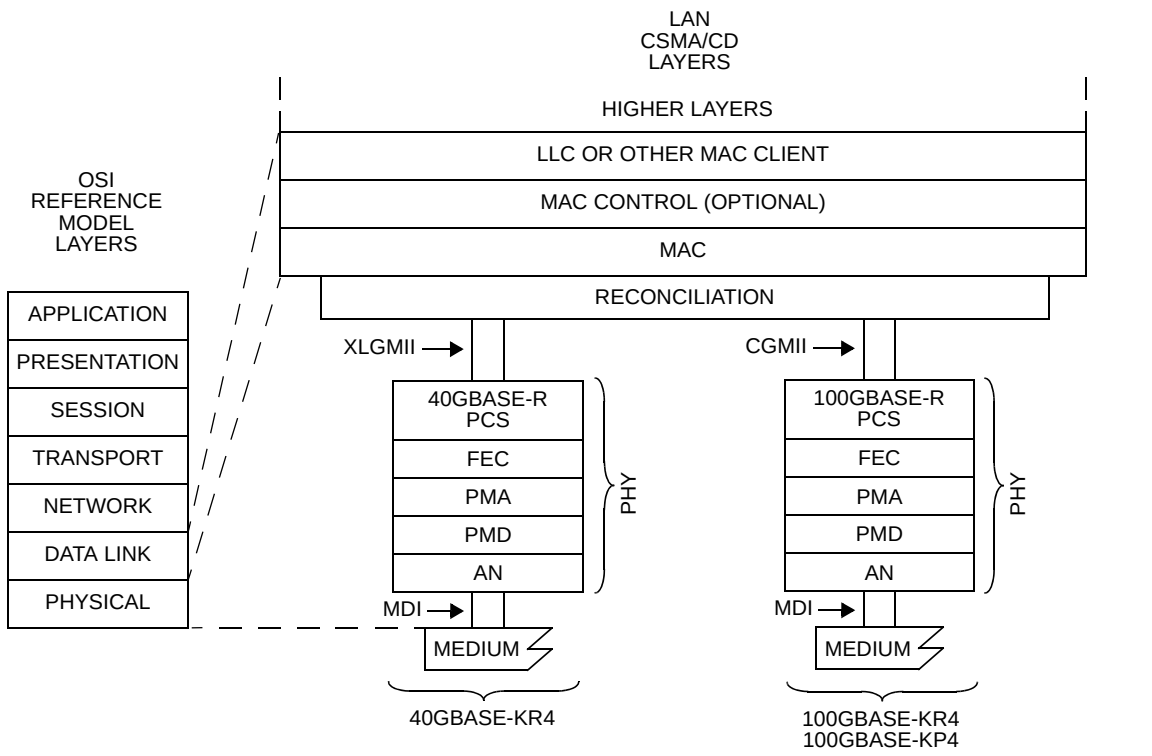
Backplane Ethernet couples the IEEE 802.3 (CSMA/CD) MAC to a family of Physical Layers defined for operation over electrical backplanes. The relationships among Backplane Ethernet, the IEEE 802.3 MAC, and the ISO Open System Interconnection (OSI) reference model are shown in Figure 69–1 and Figure 69–1a.

Change Figure 69–1 and insert Figure 69–1a as shown:



AN = AUTO-NEGOTIATION
FEC = FORWARD ERROR CORRECTION
GMII = GIGABIT 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
XGMII = 10 GIGABIT MEDIA INDEPENDENT INTERFACE

Figure 69–1—Architectural positioning of 1 Gb/s and 10 Gb/s Backplane Ethernet



AN = AUTO-NEGOTIATION
CGMII = 100 GIGABIT MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
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
XLGMII = 40 GIGABIT MEDIA INDEPENDENT INTERFACE

Figure 69–1a—Architectural positioning of 40 Gb/s and 100 Gb/s Backplane Ethernet

Delete item f) as shown:

Editor's note (to be removed prior to final publication):

It has been pointed out that Backplane Ethernet uses the MDI as an architectural reference point but it does not define any mechanical or electrical specifications for it. Apart from its presence in architectural diagrams or its occasional use as a reference point, it is not truly specified as implied by item f). See comment #280 against IEEE P802.3bj Draft 1.0.

~~f) The MDI as specified in Clause 70 for 1000BASE-KX, Clause 71 for 10GBASE-KX4, Clause 72 for 10GBASE-KR, or Clause 84 for 40GBASE-KR4.~~

69.2.1 Reconciliation sublayer and media independent interfaces

Change the first paragraph as shown:

The Clause 35 RS and GMII, the Clause 46 RS and XGMII, and the Clause 81 RS ~~and~~, XLGMII, and CGMII are employed for the same purpose in Backplane Ethernet, that being the interconnection between the MAC sublayer and the PHY.

69.2.4 Physical Layer signaling systems

Insert the following two paragraphs after the fourth paragraph:

Backplane Ethernet also specifies 100GBASE-KR4 for 100 Gb/s operation using 2-level pulse amplitude modulation (PAM) over four differential signal pairs in each direction for a total of eight pairs where the insertion loss of each pair does not exceed 35 dB at 12.9 GHz. The embodiment of 100GBASE-KR4 employs the PCS defined in Clause 82, the RS-FEC defined in Clause 91, the PMA defined in Clause 83, and the PMD defined in Clause 93.

Backplane Ethernet also specifies 100GBASE-KP4 for 100 Gb/s operation using 4-level PAM over four differential signal pairs in each direction for a total of eight pairs where the insertion loss of each pair does not exceed 33 dB at 7.0 GHz. The embodiment of 100GBASE-KP4 employs the PCS defined in Clause 82, the RS-FEC defined in Clause 91, and the PMA and PMD defined in Clause 94.

Change the last paragraph as shown:

Table 69–1 and Table 69–1a specify~~specifies~~ the correlation between nomenclature and clauses. A complete implementation conforming to one or more nomenclatures meets the requirements of the corresponding clauses.

Replace Table 69–1 and insert Table 69–1a as shown:

Table 69–1—Nomenclature and clause correlation for 1 Gb/s and 10 Gb/s Backplane Ethernet Physical Layers

Nomenclature	Clause													
	35		36	46		48	49	51	70	71	72	73	74	78
	RS	GMII	1000BASE-X PCS/PMA	RS	XGMII	10GBASE-X PCS/PMA	10GBASE-R PCS	Serial PMA	1000BASE-KX PMD	10GBASE-KX4 PMD	10GBASE-KR PMD	Auto-Negotiation	BASE-R FEC	Energy-Efficient Ethernet (EEE)
1000BASE-KX	M[a]	O[a]	M						M			M		O
10GBASE-KX4				M	O	M				M		M		O
10GBASE-KR				M	O		M	M			M	M	O	O

[a]O = Optional, M = Mandatory

Table 69–1a—Nomenclature and clause correlation for 40 Gb/s and 100 Gb/s Backplane Ethernet Physical Layers

Nomenclature	Clause															
	73	74	78	81			82		83		83A		84	91	93	94
	Auto-Negotiation	BASE-R FEC	Energy-Efficient Ethernet (EEE)	RS	XLGMII	CGMII	40GBASE-R PCS	100GBASE-R PCS	40GBASE-R PMA	100GBASE-R PMA	XLAUI	CAUI	40GBASE-KR4 PMD	100GBASE-R RS-FEC	100GBASE-KR4 PMD	100GBASE-KP4 PMA/PMD
40GBASE-KR4	M[a]	O[a]	O	M	O		M		M		O		M			
100GBASE-KR4	M		O	M		O		M		M		O		M	M	
100GBASE-KP4	M		O	M		O		M				O		M		M

[a]O = Optional, M = Mandatory

69.2.6 Low-Power Idle

Change the first sentence as shown:

With the optional EEE feature, described in Clause 78, Backplane Ethernet PHYs ~~for 10Gb/s or lower~~ can achieve lower power consumption during periods of low link utilization.

69.3 Delay constraints

Insert the following two paragraphs after the last paragraph:

For 100GBASE-KR4, normative delay specifications may be found in 81.1.4, 82.5, 83.5.4, 91.4, and 93.4 and also referenced in 80.4.

For 100GBASE-KP4, normative delay specifications may be found in 81.1.4, 82.5, 91.4, and 94.3.3 and also referenced in 80.4.

69.5 Protocol implementation conformance statement (PICS) proforma

Change the first paragraph as shown:

The supplier of a protocol implementation that is claimed to conform to any part of IEEE Std 802.3~~, Clause 70 through Clause 74,~~ demonstrates compliance by completing a protocol implementation conformance statement (PICS) proforma.

73. Auto-Negotiation for backplane and copper cable assembly

Change the first paragraph of Clause 73 as shown:

Note that although the Auto-Negotiation defined in this clause was originally intended for use with Backplane Ethernet PHYs, it is also specified for use with 40GBASE-CR4, ~~and~~ 100GBASE-CR10, and 100GBASE-CR4 PHYs.

73.3 Functional specifications

Change the last sentence of the third paragraph as shown:

These functions shall comply with the state diagrams from Figure 73–9 through Figure 73–11. The Auto-Negotiation functions shall interact with the technology-dependent PHYs through the Technology-Dependent interface (see 73.9). Technology-Dependent PHYs include 1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 40GBASE-KR4, 40GBASE-CR4, ~~and~~ 100GBASE-CR10, 100GBASE-KR4, 100GBASE-KP4, and 100GBASE-CR4.

73.5.1 DME electrical specifications

Change the last paragraph as shown:

For any multi-lane PHY, DME pages shall be transmitted only on lane 0. The transmitters on other lanes should be disabled as specified in 71.6.7, 84.7.7, ~~or~~ 85.7.7, 92.7.6, 93.7.7, or 94.3.6.6.

73.6.4 Technology Ability Field

Change Table 73–4 as shown:

Table 73–4—Technology Ability Field encoding

Bit	Technology
A0	1000BASE-KX
A1	10GBASE-KX4
A2	10GBASE-KR
A3	40GBASE-KR4
A4	40GBASE-CR4
A5	100GBASE-CR10
A6	100GBASE-KR4
A7	100GBASE-KP4
A8	100GBASE-CR4
~~A6~~A9 through A24	Reserved for future technology

Replace the second to last paragraph ("40GBASE-CR4 and...") with the following:

A PHY for operation over an electrical backplane (e.g. 1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 40GBASE-KR4, 100GBASE-KR4, 100GBASE-KP4) shall not be advertised simultaneously with a PHY for operation over a copper cable assembly (e.g. 40GBASE-CR4, 100GBASE-CR10, 100GBASE-CR4) as the MDI and physical medium are different.

Change the last paragraph as shown:

The fields A[24:~~6~~9] are reserved for future use. Reserved fields shall be sent as zero and ignored on receive.

73.7 Receive function requirements

Change the last sentence as shown:

The receive function incorporates a receive switch to control connection to the 1000BASE-KX, 10GBASE-KX4, 10GBASE-KR 40GBASE-KR4, 40GBASE-CR4,~~ or~~ 100GBASE-CR10, 100GBASE-KR4, 100GBASE-KP4, or 100GBASE-CR4 PHYs.

73.7.1 DME page reception

Change the first sentence as shown:

To be able to detect the DME bits, the receiver should have the capability to receive DME signals sent with the electrical specifications of the PHY (1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 40GBASE-KR4, 40GBASE-CR4~~, or~~ 100GBASE-CR10, 100GBASE-KR4, 100GBASE-KP4, or 100GBASE-CR4). The DME transmit signal level and receive sensitivity are specified in 73.5.1.

73.7.2 Receive Switch function

Change the last paragraph as shown:

During Auto-Negotiation, the Receive Switch function shall connect the MDI to the DME page receiver controlled by the Receive state diagram~~ to the MDI and the Receive Switch function shall also connect~~ and to the receive path of the 1000BASE-KX, 10GBASE-KX4, 10GBASE-KR, 40GBASE-KR4, 40GBASE-CR4~~, and~~ 100GBASE-CR10, 100GBASE-KR4, 100GBASE-KP4, and 100GBASE-CR4 ~~PMA receivers to the MDI~~ if the ~~PMAs are~~PHY is present.

73.7.6 Priority Resolution function

Change Table 73–5 as shown:

Table 73–5—Priority Resolution

Priority	Technology	Capability
1	100GBASE-CR4	100 Gb/s 4 lane, highest priority
2	100GBASE-KR4	100 Gb/s 4 lane
3	100GBASE-KP4	100 Gb/s 4 lane
~~1~~4	100GBASE-CR10	100 Gb/s 10 lane~~, highest priority~~
~~2~~5	40GBASE-CR4	40 Gb/s 4 lane
~~3~~6	40GBASE-KR4	40 Gb/s 4 lane
~~4~~7	10GBASE-KR	10 Gb/s 1 lane
~~5~~8	10GBASE-KX4	10 Gb/s 4 lane
~~6~~9	1000BASE-KX	1 Gb/s 1 lane, lowest priority

73.10.7 State diagram variables

Insert new values for the variable "x" as shown:

A variable with "_[x]" appended to the end of the variable name indicates a variable or set of variables as defined by "x". "x" may be as follows:

all;	represents all specific technology-dependent PMAs supported in the local device.
1GKX;	represents that the 1000BASE-KX PMA is the signal source.
10GKR;	represents that the 10GBASE-KR PMA is the signal source.
10GKX4;	represents that the 10GBASE-KX4 or 10GBASE-CX4 PMA is the signal source.
40GKR4;	represents that the 40GBASE-KR4 PMD is the signal source.
40GCR4;	represents that the 40GBASE-CR4 PMD is the signal source.
100GCR10;	represents that the 100GBASE-CR10 PMD is the signal source.
100GKR4;	represents that the 100GBASE-KR4 PMD is the signal source.
100GKP4;	represents that the 100GBASE-KP4 PMD is the signal source.
100GCR4;	represents that the 100GBASE-CR4 PMD is the signal source.
HCD;	represents the single technology-dependent PMA chosen by Auto-Negotiation as the highest common denominator technology through the Priority Resolution or parallel detection function.
notHCD;	represents all technology-dependent PMAs not chosen by Auto-Negotiation as the highest common denominator technology through the Priority Resolution or parallel detection function.
PD;	represents all of the following that are present: 1000BASE-KX PMA, 10GBASE-KX4 PMA or 10GBASE-CX4 PMA, 10GBASE-KR PMA, 40GBASE-KR4 PMD, 40GBASE-CR4 PMD, and 100GBASE-CR10 PMD.

Change the definition of the variable single_link_ready as shown:

single_link_ready

Status indicating that an_receive_idle = true and only one the of the following indications is being received:

1) link_status_[1GKX] = OK
2) link_status_[10GKX4] = OK
3) link_status_[10GKR] = OK
4) link_status_[40GKR4] = OK
5) link_status_[40GCR4] = OK
6) link_status_[100GCR10] = OK
7) link_status_[100GKR4] = OK
8) link_status_[100GKP4] = OK
9) link_status_[100GCR4] = OK

Values: false; either zero or more than one of the above indications are true or an_receive_idle = false.
true; Exactly one of the above indications is true and an_receive_idle = true.

NOTE—This variable is set by this variable definition; it is not set explicitly in the state diagrams.

73.11 Protocol implementation conformance statement (PICS) proforma for Clause 73, Auto-Negotiation for backplane and copper cable assembly[2]

Editor's note (to be removed prior to final publication):

The PICS proforma will be updated when the content of this clause stabilizes.

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

78. Energy-Efficient Ethernet

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

Change 78.1 to add 100 Gb/s Ethernet:

78.1 Overview

The optional EEE capability combines the IEEE 802.3 Media Access Control (MAC) Sublayer with a family of Physical Layers defined to support operation in the Low Power Idle (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 in to and out of the lower level of power consumption is kept small enough to be transparent to upper layer protocols and applications.

For operation over twisted-pair cabling systems, EEE ~~supports~~ may be supported by the 100BASE-TX PHY, the 1000BASE-T PHY, and the 10GBASE-T PHY. For operation over twinax cable, EEE may be supported by the 100GBASE-CR10 and the 100GBASE-CR4 PHY. For operation over electrical backplanes, EEE ~~supports~~ may be supported by the 1000BASE-KX PHY, the 10GBASE-KX4 PHY, ~~and~~ the 10GBASE-KR PHY, the 100GBASE-KR4 PHY, and the 100GBASE-KP4 PHY. EEE also ~~supports~~ may be supported by XGMII extension using the XGXS for 10 Gb/s PHYs and inter-sublayer service interfaces using the XLAUI for 40 Gb/s PHYs and CAUI for 100 Gb/s PHYs.

In addition to the above, EEE defines a 10 Mb/s MAU (10BASE-Te) with reduced transmit amplitude requirements. The 10BASE-Te MAU is fully interoperable with 10BASE-T MAUs over 100 m of class D (Category 5) or better cabling as specified in ISO/IEC 11801:1995. These requirements can also be met by Category 5 cable and components as specified in ANSI/TIA/EIA-568-B-1995. The definition of 10BASE-Te allows a reduction in power consumption.

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. Clause 78 provides an overview of EEE operation. PICS for the optional EEE capability for each specific PHY type are specified in the respective PHY clauses. Normative requirements for Data Link Layer capabilities are contained in 78.4.

Editors' Note (to be removed prior to publication):

The PHY list in 78.1.3.3.1 is not changed as this is presented with "e.g." and does not need to be definitive.

78.1.4 EEE Supported PHY types

Change the table title and body as shown in Table 78–1 for 40 Gb/s and 100 Gb/s Ethernet:

Table 78–1—Clauses associated with each ~~interface~~ PHY type

PHY type	Clause
10BASE-Te	14
100BASE-TX	24, 25
1000BASE-T	40
XGXS (XAUI)	47
1000BASE-KX	70, 35
10GBASE-T	55
10GBASE-KX4	71, 48
10GBASE-KR	72, 51, 49, 74
40GBASE-CR4	82, 83, 85, 74
40GBASE-KR4	82, 83, 84, 74
100GBASE-CR10	82, 83, 85, 74
100GBASE-CR4	82, 83, 91, 92
100GBASE-KR4	82, 83, 91, 93
100GBASE-KP4	82, 91, 94
XLAUI/CAUI	83A

78.2 LPI mode timing parameters description

Change column heading and add rows to Table 78–2 to for 100 Gb/s Ethernet:

Table 78–2 summarizes three key EEE parameters (T_s, T_q, and T_r) for supported PHYs.

78.5 Communication link access latency

Add the following to the end of 78.5

For PHYs with an operating speed of 100 Gb/s (that implement EEE) two modes of LPI operation are supported. Normal wake refers to the mode for which the transmitter ceases transmission during the quiet state (as shown in Figure 78–3); Fast wake refers to the mode for which the transmitter continues to transmit signals during the fast wake state (between the sleep and wake states) so that the receiver can resume operation with a shorter wake time.

Add rows to Table 78-4 to for 100 Gb/s Ethernet:

Add 78.5.2 for PHY extension:

Table 78–2—Summary of the key EEE parameters for supported PHY

~~Protocol~~ PHY Type	T_s (μs)		T_q (μs)		T_r (μs)	
	Min	**Max**	**Min**	**Max**	**Min**	**Max**
40GBASE-CR4	TBD	TBD	TBD	TBD	TBD	TBD
40GBASE-KR4	TBD	TBD	TBD	TBD	TBD	TBD
100GBASE-CR10	TBD	TBD	TBD	TBD	TBD	TBD
100GBASE-CR4	TBD	TBD	TBD	TBD	TBD	TBD
100GBASE-KR4	TBD	TBD	TBD	TBD	TBD	TBD
100GBASE-KP4	TBD	TBD	TBD	TBD	TBD	TBD

Editor's note (to be removed prior to final publication):

If the definitions for FEC allow more than 1 wake scenario for any PHY then multiple cases will be defined (as for 10GBASE-KR) and the text of 78.5 will be altered accordingly.

Table 78–4—Summary of the LPI timing parameters for supported PHYs

PHY Type	Case	$T_{w\_sys\_tx}$ (min) (μs)	T_{w_phy} (min) (μs)	$T_{phy\_shrink\_tx}$ (max) (μs)	$T_{phy\_shrink\_rx}$ (max) (μs)	$T_{w\_sys\_rx}$ (min) (μs)
100GBASE-CR10	Normal	TBD	TBD	TBD	TBD	TBD
	Fast Wake	TBD	TBD	TBD	TBD	TBD
100GBASE-CR4	Normal	TBD	TBD	TBD	TBD	TBD
	Fast Wake	TBD	TBD	TBD	TBD	TBD
100GBASE-KR4	Normal	TBD	TBD	TBD	TBD	TBD
	Fast Wake	TBD	TBD	TBD	TBD	TBD
100GBASE-KP4	Normal	TBD	TBD	TBD	TBD	TBD
	Fast Wake	TBD	TBD	TBD	TBD	TBD
CAUI		TBD	TBD	TBD	TBD	TBD

78.5.2 100 Gb/s PHY extension using CAUI

100 Gb/s PHYs may be extended using CAUI as a physical instantiation of the inter-sublayer service interface to separate functions between devices. The LPI signaling can operate across CAUI with no change to the PHY timing parameters described in Table 78–4 or the operation of the Data Link Layer Capabilities negotiation described in 78.4.

If the PMA Egress AUI Stop Enable (PEASE) bit (1.n.n) is asserted for any of the PMA sublayers, the PMA may stop signaling on the CAUI in the transmit direction to conserve energy. If the PEASE bit is asserted,

the RS defers sending data following deassertion of LPI by an additional time equal to $T_{w\_sys\_tx}$–$T_{w\_sys\_rx}$ as shown in Table 78–4 for each PMA with PEASE asserted (see 81.3a.2.1).

If the PMA Ingress AUI Stop Enable (PIASE) bit (1.n.n) is asserted for any of the PMA sublayers, the PMA may stop signaling on the CAUI in the receive direction to conserve energy. The receiver should negotiate an additional time for the remote T_{w_sys} (equal to $T_{w\_sys\_tx}$–$T_{w\_sys\_rx}$ for the CAUI as shown in Table 78–4) for each PMA with PIASE to be asserted before setting the PIASE bits.

78.6 Protocol implementation conformance statement (PICS) proforma for EEE Data Link Layer Capabilities[3]

Editor's Note (to be removed prior to publication):

No changes to the PICS are required for 100 Gb/s Ethernet.

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

80. Introduction to 40 Gb/s and 100 Gb/s networks

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

Delete the entire section 80.1.2 in the base document.

~~**80.1.2 Objectives**~~

~~The following are the objectives of 40 Gigabit and 100 Gigabit Ethernet:~~

- ~~a) Support the full duplex Ethernet MAC.~~
- ~~b) Preserve the IEEE 802.3 Ethernet frame format utilizing the IEEE 802.3 MAC.~~
- ~~c) Preserve minimum and maximum frame size of IEEE Std 802.3.~~
- ~~d) Support a BER better than or equal to 10^{-12} at the MAC/PLS service interface.~~
- ~~e) Provide appropriate support for Optical Transport Network (OTN).~~
- ~~f) Support a MAC data rate of 40 Gb/s.~~
- ~~g) Provide Physical Layer specifications that support 40 Gb/s operation over up to the following:~~
  - ~~1) At least 10 km on single-mode fiber (SMF)~~
  - ~~2) At least 2 km on single-mode fiber (SMF)~~
  - ~~3) At least 100 m on OM3 multimode fiber (MMF)~~
  - ~~4) At least 7 m over a copper cable assembly~~
  - ~~5) At least 1 m over a backplane~~
- ~~h) Support a MAC data rate of 100 Gb/s.~~
- ~~i) Provide Physical Layer specifications that support 100 Gb/s operation over up to the following:~~
  - ~~1) At least 40 km on single-mode fiber (SMF)~~
  - ~~2) At least 10 km on single-mode fiber (SMF)~~
  - ~~3) At least 100 m on OM3 multimode fiber (MMF)~~
  - ~~4) At least 7 m over a copper cable assembly~~

80.1.3 Relationship of 40 Gigabit and 100 Gigabit Ethernet to the ISO OSI reference model

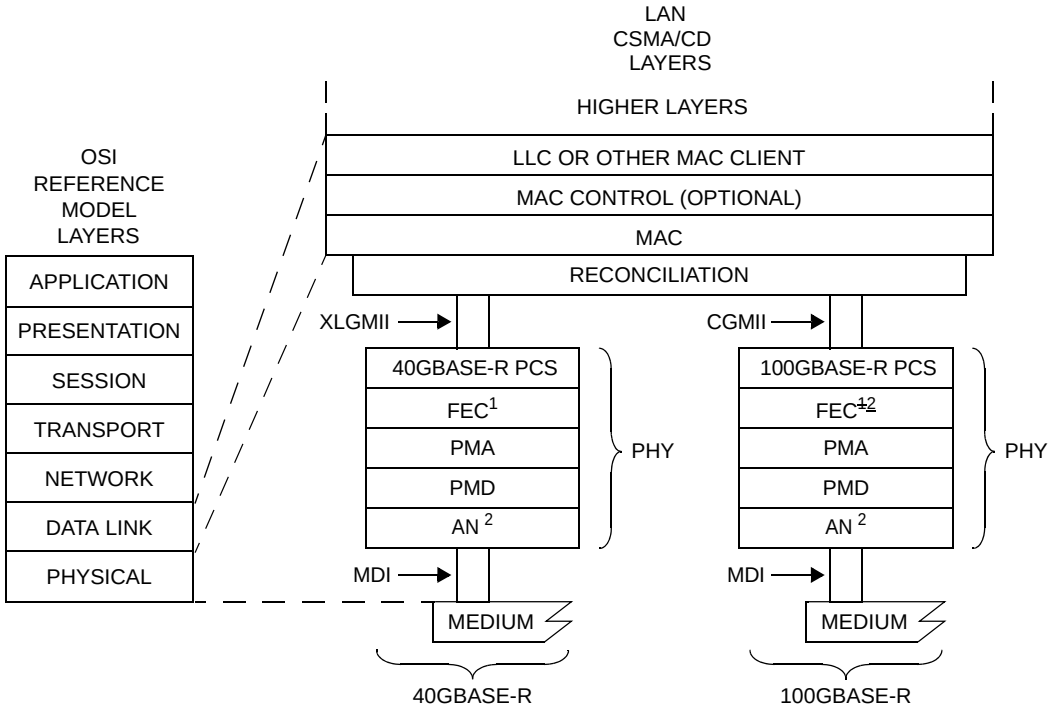
Change note h) as shown.

h) The MDIs as specified in Clause 84 for 40GBASE-KR4, in Clause 85 for 40GBASE-CR4, in Clause 86 for 40GBASE-SR4, in Clause 87 for 40GBASE-LR4, ~~and~~ in Clause 88 for 100GBASE-LR4 and 100GBASE-ER4, and in Clause 92 for 100GBASE-CR4 all use a 4 lane data path.

Add note j) as shown.

j) There is no electrical or mechanical specification of the MDI for backplane Physical Layers.

Change figure 80-1 as shown.



AN = AUTO-NEGOTIATION
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
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
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

NOTE 1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE
NOTE 2—CONDITIONAL BASED ON PHY TYPE

Figure 80–1—Architectural positioning of 40 Gigabit and 100 Gigabit Ethernet

Change 80.1.4 as shown.

80.1.4 Nomenclature

The nomenclature employed by the 40 Gigabit and 100 Gigabit Physical Layers is explained as follows.

The alpha-numeric prefix 40GBASE in the port type (e.g., 40GBASE-R) represents a family of Physical Layer devices operating at a speed of 40 Gb/s. The alpha-numeric prefix 100GBASE in the port type (e.g., 100GBASE-R) represents a family of Physical Layer devices operating at a speed of 100 Gb/s.

40GBASE-R or 100GBASE-R represents a family of Physical Layer devices using a physical coding sublayer for 40 Gb/s or 100 Gb/s operation over multiple PCS lanes based on 2-level pulse amplitude modulation (PAM) and 64B/66B block encoding (see Clause 82).

100GBASE-P represents Physical Layer devices using a physical coding sublayer for 100 Gb/s operation over multiple PCS lanes based on multi-level pulse amplitude modulation (PAM) and 64B/66B block encoding (see Clause 82).

Physical Layer devices listed in Table 80–1 are defined for operation at 40 Gb/s and 100 Gb/s.

Add the following rows to Table 80-1.

Table 80–1—40 Gb/s and 100 Gb/s PHYs

Name	Description
100GBASE-CR4	100 Gb/s PHY using 100GBASE-R encoding and Clause 91 RS-FEC over four lanes of shielded balanced copper cabling, with reach up to at least 5 m (see Clause 92)
100GBASE-KP4	100 Gb/s PHY using 100GBASE-P encoding, Clause 91 RS-FEC and 4-level pulse amplitude modulation over four lanes of an electrical back-plane, with a total insertion loss up to 33dB at 7.0 GHz (see Clause 94)
100GBASE-KR4	100 Gb/s PHY using 100GBASE-R encoding, Clause 91 RS-FEC and 2-level pulse amplitude modulation over four lanes of an electrical back-plane, with a total insertion loss up to 35dB at 12.9 GHz (see Clause 93)

80.1.5 Physical Layer signaling systems

Replace Table 80-2 with Table 80-2 and Table 80-2a as shown

Table 80–2—Nomenclature and clause correlation

~~Nomenclature~~	~~Clause~~[a]																							
	~~73~~	~~74~~	~~81~~			~~82~~		~~83~~		~~83A~~		~~83B~~		~~84~~	~~85~~		~~86~~		~~86A~~		~~87~~	~~88~~		~~89~~
	~~Auto-Negotiation~~	~~BASE-R FEC~~	~~RS~~	~~XLGMII~~	~~CGMII~~	~~40GBASE-R PCS~~	~~100GBASE-R PCS~~	~~40GBASE-R PMA~~	~~100GBASE-R PMA~~	~~XLAUI~~	~~CAUI~~	~~XLAUI~~	~~CAUI~~	~~40GBASE-KR4 PMD~~	~~40GBASE-CR4 PMD~~	~~100GBASE-CR10 PMD~~	~~40GBASE-SR4 PMD~~	~~100GBASE-SR10 PMD~~	~~XLPPI~~	~~CPPI~~	~~40GBASE-LR4 PMD~~	~~100GBASE-LR4 PMD~~	~~100GBASE-ER4 PMD~~	~~40GBASE-FR PMD~~
~~40GBASE-KR4~~	~~M~~	~~O~~	~~M~~	~~O~~		~~M~~		~~M~~		~~O~~				~~M~~										
~~40GBASE-CR4~~	~~M~~	~~O~~	~~M~~	~~O~~		~~M~~		~~M~~		~~O~~					~~M~~									
~~40GBASE-SR4~~			~~M~~	~~O~~		~~M~~		~~M~~		~~O~~		~~O~~					~~M~~		~~O~~					
~~40GBASE-FR~~			~~M~~	~~O~~		~~M~~		~~M~~		~~O~~		~~O~~												~~M~~
~~40GBASE-LR4~~			~~M~~	~~O~~		~~M~~		~~M~~		~~O~~		~~O~~							~~O~~		~~M~~			
~~100GBASE-CR10~~	~~M~~	~~O~~	~~M~~		~~O~~		~~M~~		~~M~~		~~O~~					~~M~~								
~~100GBASE-SR10~~			~~M~~		~~O~~		~~M~~		~~M~~		~~O~~		~~O~~					~~M~~		~~O~~				
~~100GBASE-LR4~~			~~M~~		~~O~~		~~M~~		~~M~~		~~O~~		~~O~~									~~M~~		
~~100GBASE-ER4~~			~~M~~		~~O~~		~~M~~		~~M~~		~~O~~		~~O~~										~~M~~	

[a]~~O = Optional, M = Mandatory.~~

Table 80–2—Nomenclature and clause correlation (40GBASE)

Nomenclature	Clause[a]														
	73	74	81		82	83	83A	83B	84	85	86	86A	87	89	78
	Auto-Negotiation	BASE-R FEC	RS	XLGMII	40GBASE-R PCS	40GBASE-R PMA	XLAUI	XLAUI	40GBASE-KR4 PMD	40GBASE-CR4 PMD	40GBASE-SR4 PMD	XLPPI	40GBASE-LR4 PMD	40GBASE-FR PMD	EEE
40GBASE-KR4	M	O	M	O	M	M	O		M						O
40GBASE-CR4	M	O	M	O	M	M	O			M					O
40GBASE-SR4			M	O	M	M	O	O			M	O			
40GBASE-FR			M	O	M	M	O	O						M	
40GBASE-LR4			M	O	M	M	O	O				O	M		

[a]O = Optional, M = Mandatory.

Table 80–2a—Nomenclature and clause correlation (100GBASE)

Nomenclature	Clause[a]																	
	73	74	81		82	83	83A	83B	85	86	86A	88		78	91	92	93	94
	Auto-Negotiation	BASE-R FEC	RS	CGMII	100GBASE-R PCS	100GBASE-R PMA	CAUI	CAUI	100GBASE-CR10 PMD	100GBASE-SR10 PMD	CPPI	100GBASE-LR4 PMD	100GBASE-ER4 PMD	EEE	BASE-R RS FEC	100GBASE-CR4 PMD	100GBASE-KR4 PMD	100GBASE-KP4 PMD
100GBASE-CR10	M	O	M	O	M	M	O		M					O				
100GBASE-CR4	M		M	O	M	M		O						O	M	M		
100GBASE-KR4	M		M	O	M	M		O						O	M		M	
100GBASE-KP4	M		M	O	M									O	M			M
100GBASE-SR10			M	O	M	M	O	O		M	O							
100GBASE-LR4			M	O	M	M	O	O				M						
100GBASE-ER4			M	O	M	M	O	O					M					

[a]O = Optional, M = Mandatory.

Change 80.2.3 as shown.

80.2.3 Forward Error Correction (FEC) sublayers

~~The~~ A Forward Error Correction sublayer is an optional sublayer for most 40GBASE-R and 100GBASE-R copper and backplane PHYs, and is mandatory for certain PHYs. The FEC sublayer can be placed in between the PCS and PMA sublayers or between two PMA sublayers~~, is instantiated for each PCS lane, and operates autonomously on a per PCS lane basis~~.

The BASE-R FEC ~~sublayer~~ (see Clause 74) is instantiated for each PCS lane, and operates autonomously on a per PCS lane basis ~~specified in Clause 74~~. The Reed-Solomon FEC (see Clause 91) is instantiated once and requires 20 PCS lanes and 4 PMA lanes for operation.

80.2.4 Physical Medium Attachment (PMA) sublayer

Change the second paragraph of 80.2.4 as shown.

The 40GBASE-R and 100GBASE-R PMAs are specified in Clause 83; the PMA for 100GBASE-KP4 is specified in Clause 94.

80.2.5 Physical Medium Dependent (PMD) sublayer

Change the second paragraph of 80.2.5 as shown.

The 40GBASE-R and 100GBASE-R PMDs and their corresponding media are specified in Clause 84 through Clause 89 and Clause 92 through Clause 94.

80.2.6 Auto-Negotiation

Change the last sentence as shown.

Clause 73 Auto-Negotiation is used by the 40 Gb/s and 100 Gb/s backplane PHYs (40GBASE-KR4, 100GBASE-KR4, and 100GBASE-KP4~~see Clause 84~~) and the 40 Gb/s and 100 Gb/s copper PHYs (40GBASE-CR4, ~~and~~ 100GBASE-CR10, and 100GBASE-CR4~~see Clause 85~~).

80.3.1 Inter-sublayer service interface

Add the following at the end of 80.3.1:

If the optional Energy Efficient Ethernet (EEE) capability is supported (see Clause 78, 78.3) then the inter-sublayer service interface includes two additional primitives defined as follows:

IS_TX_MODE.request
IS_RX_MODE.indication

The IS_TX_MODE.request primitive is used to communicate the state of the PCS LPI transmit function to other sublayers in the PHY. The IS_RX_MODE.indication primitive is used to communicate the state of the received signal to the PCS (and other sublayers of the PHY).

80.3.2 Inter-sublayer service interface

Change the second paragraph of 80.3.2 as shown:

Examples of inter-sublayer service interfaces for 40GBASE-R and 100GBASE-R with their corresponding

instance names are illustrated in Figure 80–2, ~~and~~ Figure 80–3 and Figure 80–3a. For example, the primitives for one instance of the inter-sublayer service interface, named the PMD service interface, are identified as follows:

Change Figure 80–3 and add Figure 80-3a as shown:

Add 80.3.3.4 and 80.3.3.5 as shown:

80.3.3.4 IS_TX_MODE.request

The IS_TX_MODE.request primitive communicates the tx_mode parameter generated by the PCS Transmit Process for EEE capability to invoke the appropriate PMA, FEC and PMD transmit EEE states. Without EEE capability, the primitive is never invoked and the sublayers behave as if tx_mode = DATA.

80.3.3.4.1 Semantics of the service primitive

IS_TX_MODE.request(tx_mode)

The tx_mode parameter takes on one of eight values: DATA, SLEEP, QUIET, FW, ALERT, RF_ALERT, WAKE or RF_WAKE.

80.3.3.4.2 When generated

This primitive is generated to indicate the low power mode of the transmit path.

80.3.3.4.3 Effect of receipt

The specific effect of receipt of this primitive is defined by the sublayer that receives this primitive. In general, when tx_mode is DATA the sublayer operates normally and when tx_mode is QUIET, the sublayer may go into a low power mode.

80.3.3.5 IS_RX_MODE.indicate

The IS_RX_MODE.request primitive communicates the rx_mode parameter generated by the PMA/PMD for EEE capability to reflect the state of the received signal. Without EEE capability, the primitive is never invoked and the sublayers behave as if rx_mode = DATA.

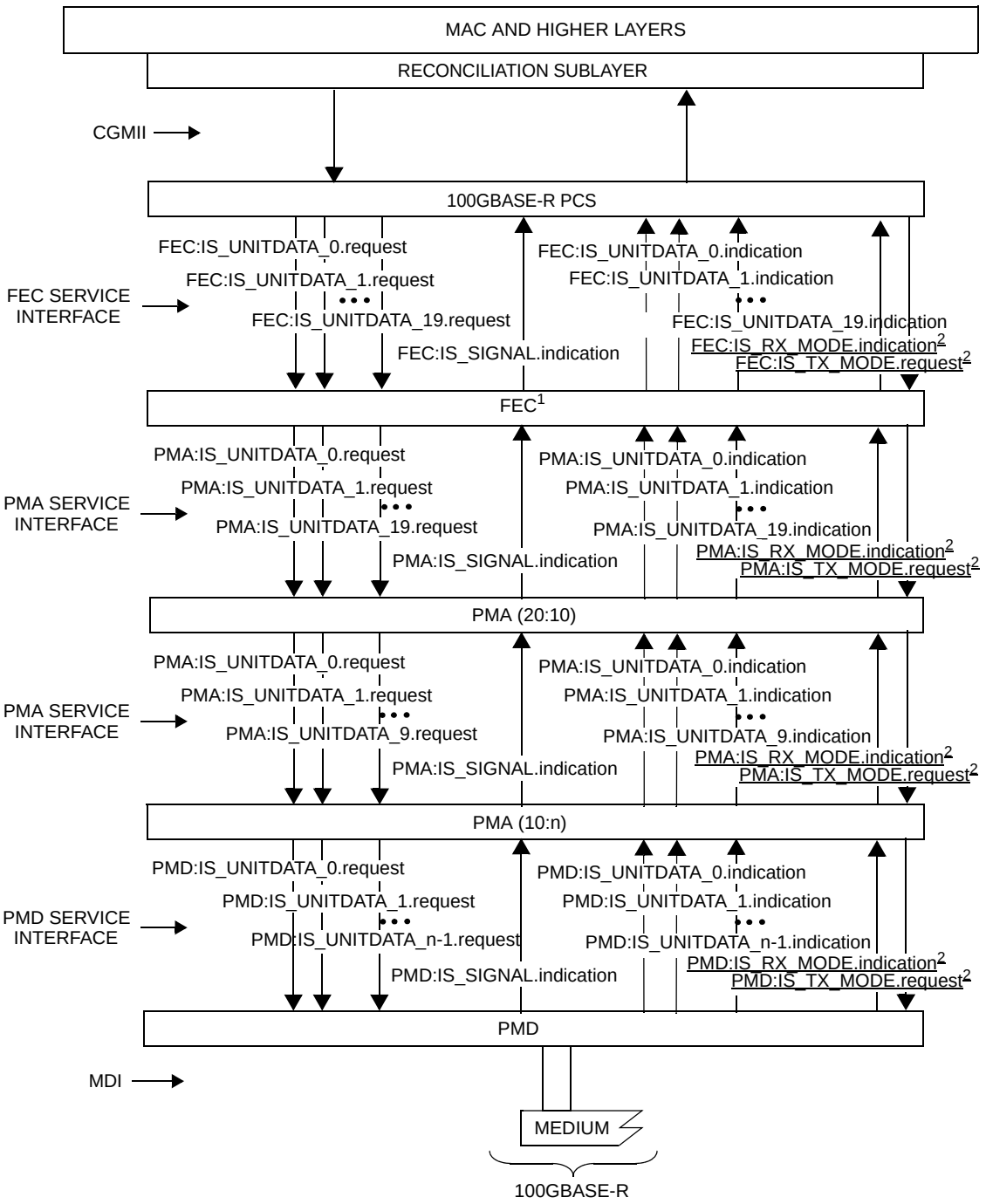
80.3.3.5.1 Semantics of the service primitive

IS_RX_MODE.indicate(rx_mode)

The rx_mode parameter takes on one of four values: DATA, QUIET, or ALERT.

80.3.3.5.2 When generated

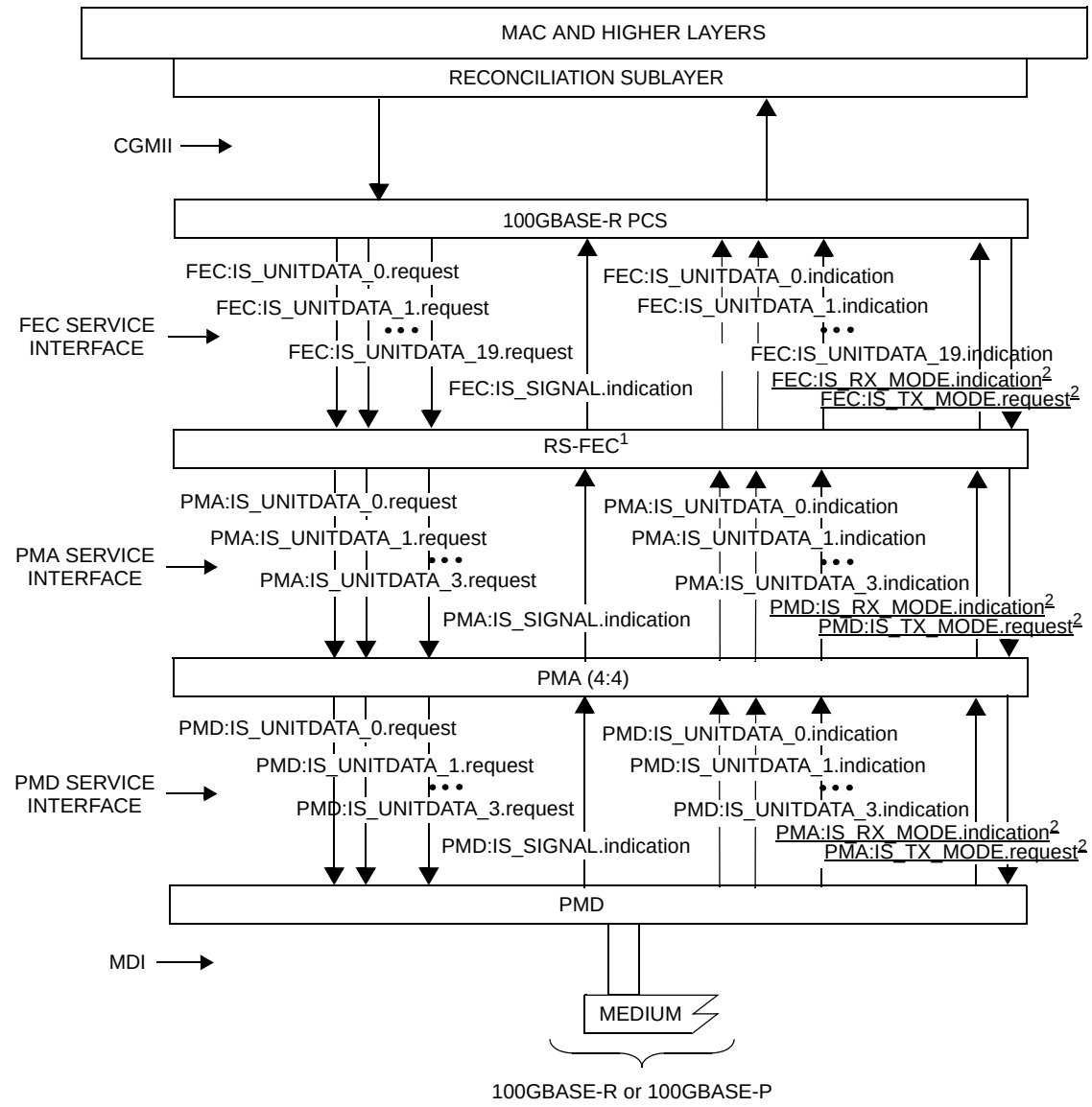
This primitive is generated to indicate the state of the received signal.



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
n = NUMBER OF PARALLEL STREAMS OF DATA UNITS

~~NOTE 1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE~~
NOTE 1—CONDITIONAL BASED ON PHY TYPE
NOTE 2—FOR OPTIONAL EEE SUPPORT

Figure 80–3—100GBASE-R inter-sublayer service interfaces



CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
RS-FEC = REED-SOLOMON FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

NOTE 1—CONDITIONAL BASED ON PHY TYPE
NOTE 2—FOR OPTIONAL EEE SUPPORT

Figure 80–3a—100GBASE-R inter-sublayer service interfaces with RS-FEC

80.3.3.5.3 Effect of receipt

The specific effect of receipt of this primitive is defined by the sublayer that receives this primitive. In general, when rx_mode is DATA the sublayer operates normally and when rx_mode is QUIET, the sublayer may go into a low power mode.

80.4 Delay constraints

Insert rows in Table 80-3 as shown:

Table 80–3—Sublayer delay constraints

Sublayer	Maximum (bit time)[a]	Maximum (pause_quanta)[b]	Maximum (ns)	Notes[c]
100GBASE-R RS-FEC	TBD	TBD	TBD	See 91.4.
100GBASE-CR4 PMD	TBD	TBD	TBD	Does not include delay of cable medium. See 92.4.
100GBASE-KR4 PMD	TBD	TBD	TBD	See 93.4.
100GBASE-KP4 PMA/PMD	TBD	TBD	TBD	See 94.2.7.

[a] Note that for 40GBASE-R, 1 bit time (BT) is equal to 25 ps and for 100GBASE-R, 1 bit time (BT) is equal to 10 ps. (See 1.4.110 for the definition of bit time.)
[b] Note that for 40GBASE-R, 1 pause_quantum is equal to 12.8 ns and for 100GBASE-R, 1 pause_quantum is equal to 5.12 ns. (See 31B.2 for the definition of pause_quanta.)
[c] Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.

80.5 Skew constraints

Change NOTE 1 in Figure 80-4 as shown:

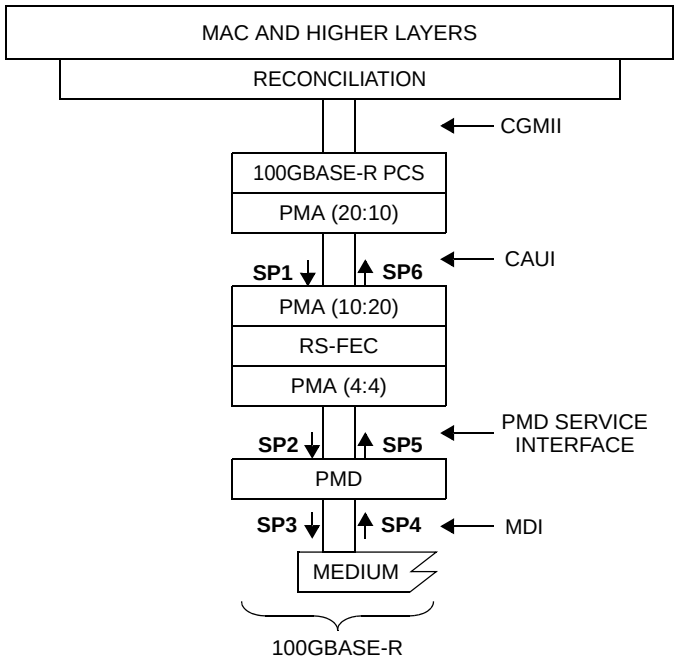
NOTE1—~~OPTIONAL OR OMITTED DEPENDING~~ CONDITIONAL BASED ON PHY TYPE

Change NOTE 1 in Figure 80-5 as shown:

NOTE1—~~OPTIONAL OR OMITTED DEPENDING~~ CONDITIONAL BASED ON PHY TYPE

Add Figure 80–5a with introductory text after Figure 80-5:

The skew points are similarly illustrated for a PHY incorporating RS-FEC (see Clause 91) in Figure 80–5a.



CAUI = 100 Gb/s 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 80–5a—100GBASE-R Skew points with RS-FEC and CAUI

Change the "Notes" columns in Table 80-4 and 80-5 as shown:

Table 80–4—Summary of Skew constraints

Skew points	Maximum Skew (ns)[a]	Maximum Skew for 40GBASE-R PCS lane (UI)[b]	Maximum Skew for 100GBASE-R PCS lane (UI)[c]	Notes[d]
SP1	29	≈ 299	≈ 150	See 83.5.3.1
SP2	43	≈ 443	≈ 222	See 83.5.3.3 or 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP3	54	≈ 557	≈ 278	See 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP4	134	≈ 1382	≈ 691	See 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP5	145	≈ 1495	≈ 748	See 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP6	160	≈ 1649	≈ 824	See 83.5.3.5
At PCS receive	180	≈ 1856	≈ 928	See 82.2.12

[a]The Skew limit includes 1 ns allowance for PCB traces that are associated with the Skew points.
[b]The symbol ≈ indicates approximate equivalent of maximum Skew in UI for 40GBASE-R, based on 1 UI equals 96.969697 ps at PCS lane signaling rate of 10.3125 GBd.
[c]The symbol ≈ indicates approximate equivalent of maximum Skew in UI for 100GBASE-R, based on 1 UI equals 193.939394 ps at PCS lane signaling rate of 5.15625 GBd.
[d]Should there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sublayer clause prevails.

Table 80–5—Summary of Skew Variation constraints

Skew points	Maximum Skew Variation (ns)	Maximum Skew Variation for 10.3125 GBd PMD lane (UI)[a]	Maximum Skew Variation for 25.78125 GBd PMD lane (UI)[b]	Notes[c]
SP1	0.2	≈ 2	N/A	See 83.5.3.1
SP2	0.4	≈ 4	≈ 10	See 83.5.3.3 or 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP3	0.6	≈ 6	≈ 15	See 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP4	3.4	≈ 35	≈ 88	See 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP5	3.6	≈ 37	≈ 93	See 83.5.3.4 or 84.5 or 85.5 or 86.3.2 or 87.3.2 or 88.3.2 or 89.3.2 or 92.5 or 93.5
SP6	3.8	≈ 39	N/A	See 83.5.3.5
At PCS receive	4	≈ 41	N/A	See 82.2.12

[a]The symbol ≈ indicates approximate equivalent of maximum Skew Variation in UI for 40GBASE-R, based on 1 UI equals 96.969697 ps at PMD lane signaling rate of 10.3125 GBd.

[b]The symbol ≈ indicates approximate equivalent of maximum Skew Variation in UI for 100GBASE-R, based on 1 UI equals 38.787879 ps at PMD lane signaling rate of 25.78125 GBd.
[c]Should there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sublayer clause prevails.

80.7 Protocol implementation conformance statement (PICS) proforma for Clause 80, Introduction to 40 Gb/s and 100 Gb/s networks[4]

Editor's note (to be removed prior to final publication):

The PICS proforma will be updated when the content of this clause stabilizes.

Change the first paragraph of 80.7 as shown:

The supplier of a protocol implementation that is claimed to conform to any part of IEEE Std 802.3~~, Clause 45, Clause 73, Clause 74, Clause 81 through Clause 89, and related annexes~~ demonstrates compliance by completing a protocol implementation conformance statement (PICS) proforma.

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

81. Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation (XLGMII and CGMII)

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

81.1 Overview

Change NOTE 1 in Figure 81-1 as shown:

NOTE1—~~OPTIONAL OR OMITTED DEPENDING~~ CONDITIONAL BASED ON PHY TYPE

81.1.1 Summary of major concepts

Insert item g) at the end of the list of major concepts:

g) The CGMII may also support Low Power Idle (LPI) signaling for PHY types supporting Energy Efficient Ethernet (EEE) (see Clause 78).

81.1.7 Mapping of XLGMII/CGMII signals to PLS service primitives

Change 81.1.7 for LPI operation:

The Reconciliation Sublayer (RS) shall map the signals provided at the XLGMII/CGMII to the PLS service primitives defined in Clause 6. The PLS service primitives provided by the RS and described here behave in exactly the same manner as defined in Clause 6. Full duplex operation only is implemented at 40 Gb/s and 100 Gb/s; therefore, PLS service primitives supporting CSMA/CD operation are not mapped through the RS to the XLGMII/CGMII.This behavior and restrictions are the same as described in 22.6a, with the details of the signaling described in 81.3. LPI_REQUEST shall not be set to ASSERT unless the attached link has been operational for at least one second (i.e. link_status = OK, according to the underlying PCS/PMA).

EEE capability requires the use of the MAC defined in Annex 4A 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 its low power state.

Mappings for the following primitives are defined for 40 Gb/s and 100 Gb/s operation:

PLS_DATA.request
PLS_DATA.indication
PLS_CARRIER.indication
PLS_SIGNAL.indication
PLS_DATA_VALID.indication

81.1.7.3 Mapping of PLS_CARRIER.indication

Change 81.1.7.3 for carrier indication definition:

40 Gb/s and 100 Gb/s operation supports full duplex operation only. The RS never generates this primitive for PHYs that do not support EEE.

For PHYs that support EEE capability, CARRIER_STATUS is set in response to LPI_REQUEST as shown in Figure 81–10a.

81.3 XLGMII/CGMII functional specifications

81.3.1.2 TXC<7:0> (transmit control)

Insert the following at the end of 81.3.1.2:

A PHY with EEE capability shall interpret the combination of TXC and TXD as shown in Table 81–3 as an assertion of LPI. Transition into and out of the LPI state is shown in Figure 81–6a.

Change Table 81–3 as follows:

Table 81–3—Permissible encodings of TXC and TXD

TXC	TXD	Description	PLS_DATA.request parameter
~~1~~	~~00 through 06~~	~~Reserved~~	~~—~~
1	00 through 05	Reserved	—
1	06	Only valid on all eight lanes simultaneously to request LPI	No applicable parameter (Normal inter-frame)

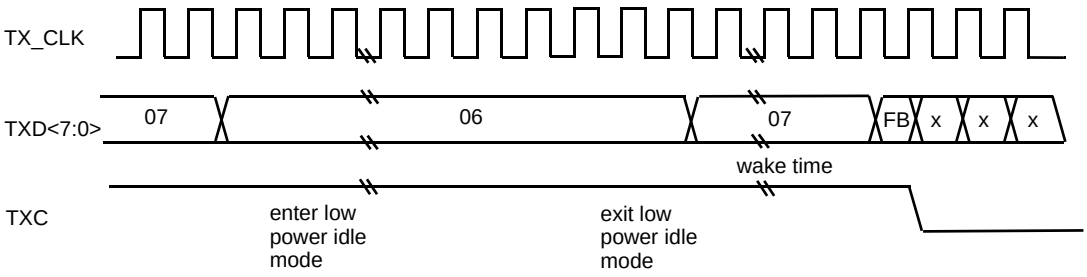
Insert 81.3.1.5 after 81.3.1.4 for transmit LPI transition:

81.3.1.5 Transmit direction LPI transition

LPI operation and the LPI client are described in 78.1. The RS requests the PHY to transition to the LPI state by asserting TXC and setting TXD to 0x06 (in all lanes). The RS maintains the same state for these signals for the entire time that the PHY is to remain in the LPI state.

The RS asserts TXC and asserts IDLE on lanes 0 to 7 in order to make the PHY transition out of the LPI state. The RS should not present a start code for valid transmit data until after the wake up time specified for the PHY. The wake times are shown in Table 78–4

Figure 81–6a shows the behavior of TXC and TXD<7:0> during the transition into and out of the LPI state.



Note: TXC and TXD are shown for one lane, all 8 lanes behave identically during LPI

Figure 81–6a—LPI transition

Table 81–3 summarizes the permissible encodings of TXD<63:0>, TXC<7:0>.

81.3.2.2 RXC<7:0> (receive control)

Insert new row and change Table 81–4 as follows:

Table 81–4—Permissible lane encodings of RXD and RXC

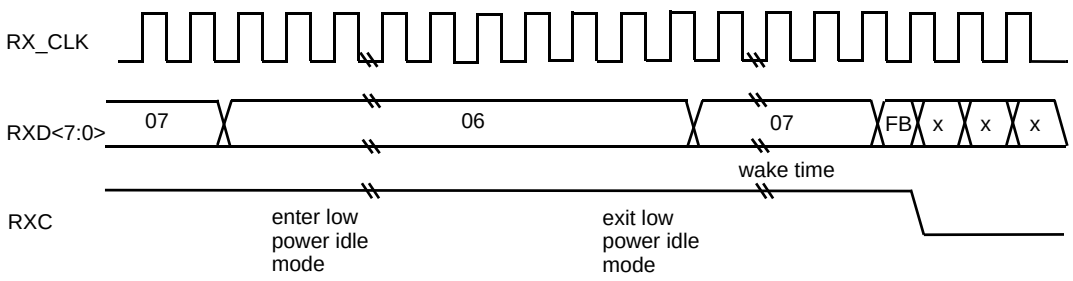
RXC	RXD	Description	PLS_DATA.indication parameter
~~1~~	~~00 through 06~~	~~Reserved~~	~~—~~
1	00 through 05	Reserved	—
1	06	Only valid on all eight lanes simultaneously to indicate LP_IDLE is asserted	No applicable parameter (Normal inter-frame)

Insert 81.3.2.4 after 81.3.2.3 for receive LPI transition:

81.3.2.4 Receive direction LPI transition

LPI operation and the LPI client are described in 78.1. When the PHY receives signals from the link partner to indicate transition into the low power state it indicates this to the RS by asserting RXC and setting RXD to 0x06 (in all lanes). The PHY maintains these signals in this state while it remains in the LPI state. When the PHY receives signals from the link partner to indicate transition out of the LPI state it indicates this to the RS by asserting RXC and asserting idle on all lanes 0 to 7 to return to a normal interframe state. The RS shall interpret the LPI coding as shown in Table 81–4.

Figure 81–8a shows the behavior of RXC and RXD<7:0> during LPI transitions.



Note: RXC and RXD are shown for one lane, all 8 lanes behave identically during LPI
Note: In some instances, LPI may be followed by characters other than IDLE during wake time

Figure 81–8a—LPI transition

81.3.4 Link fault signaling

Change 81.3.4 as follows:

Link fault signaling operates between the remote RS and the local RS. Faults detected between the remote RS and the local RS are received by the local RS as Local Fault. Only an RS originates Remote Fault signals. The behavior of the fault signaling is the same as it is for Clause 46 with the exception that the ordered sets are aligned to eight byte boundaries, padding lanes 4 to 7 with 0x00.

Clause 46 uses the term column when describing data transfers on the XGMII. The eight lanes of data and control transferred per clock cycle on XLGMII/CGMII are equivalent to a column in the following description of link fault signaling.

Sublayers within the PHY are capable of detecting faults that render a link unreliable for communication. Upon recognition of a fault condition, a PHY sublayer indicates Local Fault status on the data path. When this Local Fault status reaches an RS, the RS stops sending MAC data or LPI, and continuously generates a Remote Fault status on the transmit data path (possibly truncating a MAC frame being transmitted). When Remote Fault status is received by an RS, the RS stops sending MAC data or LPI, and continuously generates Idle control characters. When the RS no longer receives fault status messages, it returns to normal operation, sending MAC data or LPI. Note that this behavior only supports bidirectional operation.

Status is signaled in an eight byte Sequence ordered_set as shown in Table 81–5. The PHY indicates Local Fault with a Sequence control character in lane 0 and data characters of 0x00 in lanes 1, 2, 4, 5, 6, and 7 plus a data character of 0x01 in lane 3. The RS indicates a Remote Fault with a Sequence control character in lane 0 and data characters of 0x00 in lanes 1, 2, 4, 5, 6, and 7 plus a data character of 0x02 in lane 3. Though most fault detection is on the receive data path of a PHY, in some specific sublayers, faults can be detected on the transmit side of the PHY. This is also indicated by the PHY with a Local Fault status. All other values in lanes 1 to 3 not shown in Table 81–5 are reserved. The link fault signaling state diagram allows future standardization of reserved Sequence ordered sets for functions other than link fault indications.

The RS reports the fault status of the link. Local Fault indicates a fault detected on the receive data path between the remote RS and the local RS. Remote Fault indicates a fault on the transmit path between the local RS and the remote RS. The RS shall implement the link fault signaling state diagram (see Figure 81–9).

Insert a new section, 81.3a before 81.4:

81.3a LPI Assertion and Detection

Certain PHYs support Energy Efficient Ethernet (see Clause 78). PHYs with EEE capability support LPI assertion and detection. LPI operation and the LPI client are described in 78.1. LPI signaling allows the RS to signal to the PHY and to the link partner that a break in the data stream is expected and components may use this information to enter power saving modes that require additional time to resume normal operation. Similarly, it allows the LPI client to understand that the link partner has sent such an indication.

The LPI assertion and detection mechanism fits conceptually between the PLS Service Primitives and the CGMII signals as shown in Figure 81–9a.

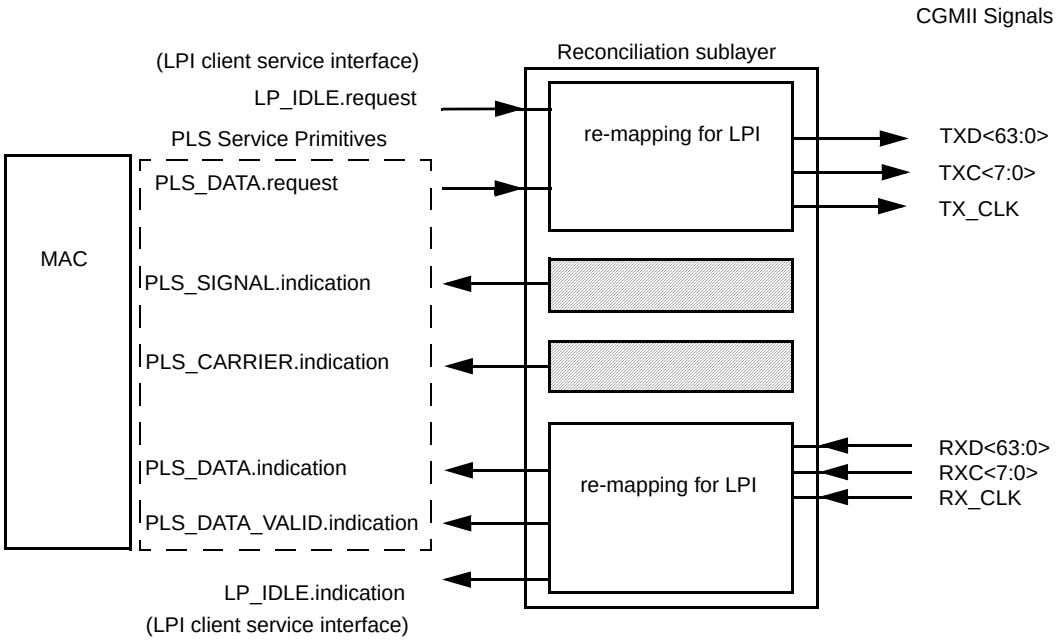


Figure 81–9a—LPI assertion and detection mechanism

The definition of TXC<7:0> and TXD<63:0> is derived from the state of PLS_DATA.request (81.1.7), except when it is overridden by an assertion of LP_IDLE.request.

Similarly, RXC<7:0> and RXD<63:0> are mapped to PLS_DATA.indication except when LP_IDLE is detected.

PLS_CARRIER.indication(CARRIER_STATUS) will be set to CARRIER_ON when the link is in LPI mode. See 81.1.7.3.

The timing of PLS_CARRIER.indication when used for the LPI function is controlled by the LPI transmit state diagram.

81.3a.1 LPI messages

LP_IDLE.indication(LPI_INDICATION)

A primitive that indicates to the LPI client that the PHY has detected the assertion or de-assertion of LPI from the link partner.

Values: DEASSERT: The link partner is operating with normal inter-frame behavior (default).
ASSERT: The link partner has asserted LPI.

LP_IDLE.request(LPI_REQUEST)

The LPI_REQUEST parameter can take one of two values: ASSERT or DE-ASSERT. ASSERT

initiates the signaling of LPI to the link partner. DE-ASSERT stops the signaling of LPI to the link partner. The effect of receipt of this primitive is undefined if link_status is not OK (see 28.2.6.1.1) or within 1 second of the change of link_status to OK.

81.3a.2 Transmit LPI state diagram

The operation of LPI in the PHY requires that the MAC does not send valid data for a time after LPI has been de-asserted as governed by resolved Transmit T_{w_sys} defined in 78.4.2.3.

This wake up time is enforced by the transmit LPI state diagram using CARRIER_SENSE.indication. The implementation shall conform to the behavior described by the transmit LPI state diagram shown in Figure 81–10a.

Editor's note (to be removed prior to final publication):

The state diagram conventions described in 80.6 apply to all of the state diagrams in this clause.

81.3a.2.1 Variables and counters

The transmit LPI state diagram uses the following variables and counters:

LPI_CARRIER_STATUS
: The LPI_CARRIER_STATUS variable indicates how the CARRIER_STATUS parameter is controlled by the LPI_REQUEST parameter. The LPI_CARRIER_STATUS is either TRUE or FALSE as determined by the Transmit LPI state diagram in Figure 81–10a.

power_on
: Condition that is true until such time as the power supply for the device that contains the RS has reached the operating region.
Values: FALSE: The device is completely powered (default).
TRUE: The device has not been completely powered.

reset
: Used by management to control the resetting of the RS.
Values: FALSE: Do not reset the RS (default).
TRUE: Reset the RS.

tw_timer
: A timer that counts, in microseconds, the time since the de-assertion of LPI. The terminal count of the timer is the value of the resolved $T_{w\_sys\_tx}$ as defined in 78.2. If PMA Ingress AUI Stop Enable (PIASE) bit (1.TBD) is asserted for any of the PMA sublayers, the terminal count of the timer is the value of the resolved $T_{w\_sys\_tx}$ as defined in 78.2 plus additional time equal to $T_{w\_sys\_tx}$–$T_{w\_sys\_rx}$ for the CAUI as shown in Table 78–4 for each PMA with PIASE to be asserted. The signal tw_timer_done is asserted when tw_timer reaches its terminal count.

81.3a.2.2 State Diagram

81.3a.3 Considerations for transmit system behavior

The transmit system should expect that egress data flow will be halted for at least resolved $T_{w\_sys\_tx}$ (see 78.2) time, in microseconds, after it requests the de-assertion of LPI. Buffering and queue management should be designed to accommodate this.

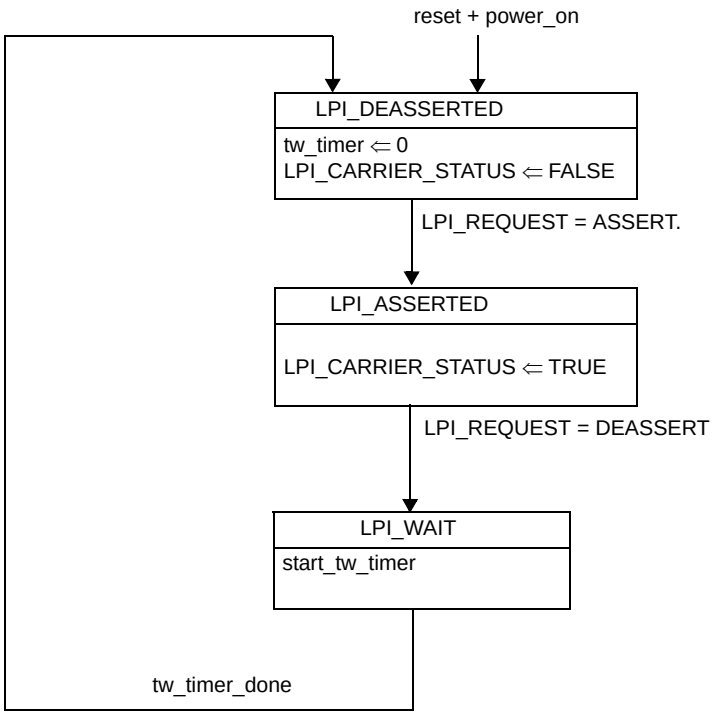


Figure 81–10a—Transmit LPI State Diagram

81.3a.3.1 Considerations for receive system behavior

The mapping function of the Reconciliation Sublayer shall continue to signal DATA_NOT_VALID on PLS_DATA_VALID.indication while it is detecting LP_IDLE on the CGMII. The receive system should be aware that data frames may arrive at the CGMII following the de-assertion of LPI_INDICATION with a delay corresponding to the link partner's resolved $T_{w\_sys\_rx}$ (as specified in 78.5) time, in microseconds.

If the PMA Ingress AUI Stop Enable (PIASE) bit (1.TBD) is asserted for any of the PMA sublayers, the PMA may stop signaling on the CAUI in the receive direction to conserve energy. The receiver should negotiate an additional time for the remote T_{w_sys} (equal to $T_{w\_sys\_tx}$–$T_{w\_sys\_rx}$ for the CAUI as shown in Table 78–4) for each PMA with PIASE to be asserted before setting the PIASE bits.

81.4 Protocol implementation conformance statement (PICS) proforma for Clause 81, Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation[5]

81.4.2.3 Major capabilities/options

Insert the following row into table 81.4.2.3:

Item	Feature	Subclause	Value/Comment	Status	Support
*LPI	Implementation of LPI	81.1.7		O	Yes [ ] No [ ]

Insert the new subclause 81.4.3.6 after 81.4.3.5 for LPI functions:

81.4.3.6 LPI functions

Item	Feature	Subclause	Value/Comment	Status	Support
L1	Assertion of LPI in Tx direction	81.3.1.2	As defined in Table 81–3	LPI:M	Yes [ ] N/A [ ]
L2	Assertion of LPI in Rx direction	81.3.2.2	As defined in Table 81–4	LPI:M	Yes [ ] N/A [ ]

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

82. Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R

Editor's note (to be removed prior to final publication):

The state diagram conventions described in 82.2.18.1 apply to all of the state diagrams in this clause.

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

82.1.3 Summary of 40GBASE-R and 100GBASE-R sublayers

Change NOTE 1 in Figure 82-1 as follows:

NOTE 1—~~OPTIONAL OR OMITTED DEPENDING~~ CONDITIONAL BASED ON PHY TYPE

Change 82.1.4 as follows:

82.1.4 Inter-sublayer interfaces

The upper interface of the PCS may connect to the Reconciliation Sublayer through the XLGMII/CGMII. The lower interface of the PCS connects to the PMA sublayer to support a PMD. If the optional FEC sublayer is implemented (see Clause 74) and an optional physical instantiation, i.e., XLAUI or CAUI, is not implemented directly below the PCS sublayer, then the lower interface connects to the FEC sublayer. For Physical Layers that use Clause 91 RS-FEC, if an optional physical instantiation, i.e. CAUI, is not implemented directly below the PCS sublayer, then the lower interface connects to the FEC sublayer. The 40GBASE-R PCS has a nominal rate at the PMA/FEC service interface of 10.3125 Gtransfers/s per PCS lane, which provides capacity for the MAC data rate of 40 Gb/s. The 100GBASE-R PCS has a nominal rate at the PMA/FEC service interface of 5.15625 Gtransfers/s per PCS lane, which provides capacity for the MAC data rate of 100 Gb/s.

It is important to note that, while this specification defines interfaces in terms of bits, octets, and frames, implementations may choose other data-path widths for implementation convenience.

82.1.4.1 PCS service interface (XLGMII/CGMII)

The PCS service interface allows the 40GBASE-R or 100GBASE-R PCS to transfer information to and from a PCS client. The PCS client is the Reconciliation Sublayer. The PCS Service Interface is precisely defined as the Media Independent Interface (XLGMII/CGMII) in Clause 81.

82.1.4.2 Physical Medium Attachment (PMA) or Forward Error Correction (FEC) service interface

The PMA or FEC service interface for the PCS is described in an abstract manner and does not imply any particular implementation. The PMA/FEC Service Interface supports the exchange of encoded data between the PCS and PMA or FEC sublayer. The PMA or FEC service interface is defined in 83.3 or 94.2.1 and is an instance of the inter-sublayer service interface definition in 80.3 or 91.2.

Change Figure 82-2 in 82.1.5 for the functional block diagram:

82.1.5 Functional block diagram

Figure 82–2 provides a functional block diagram of the 40GBASE-R PCS and 100GBASE-R PCS.

Change 82.2.3.4 for the control codes

82.2.3.4 Control codes

The same set of control characters are supported by the XLGMII/CGMII and the PCS. The representations of the control characters are the control codes. XLGMII/CGMII encodes a control character into an octet (an eight bit value). The 40GBASE-R and 100GBASE-R PCS encode the start and terminate control characters implicitly by the block type field. The 40GBASE-R and 100GBASE-R PCS encode the ordered_set control codes using the block type field. The 40GBASE-R and 100GBASE-R PCS encode each of the other control characters into a 7-bit control code.

The control characters and their mappings to 40GBASE-R and 100GBASE-R control codes and XLGMII/CGMII control codes are specified in Table 82–1. All XLGMII/CGMII, 40GBASE-R, and 100GBASE-R control code values that do not appear in the table shall not be transmitted and shall be considered an error if received. The ability to transmit or receive Low Power Idle (LPI) is required for PHYs that support EEE (see Clause 78). If EEE is not supported LPI shall not be transmitted and shall be treated as an error if received.

.Insert row in Table 82-1 for LPI coding:

Table 82–1—Control codes

Control character	Notation	XLGMII/ CGMII control code	40/100GBASE-R O code	40GBASE-R and 100GBASE-R control code
idle	/I/	0x07		0x00
LPI	/LI/	0x06		0x06
start	/S/	0xFB		Encoded by block type field

Change 82.2.3.6 for LPI definition:

82.2.3.6 Idle (/I/)

Idle control characters (/I/) are transmitted when idle control characters are received from the XLGMII/CGMII. Idle control characters may be added or deleted by the PCS to adapt between clock rates.

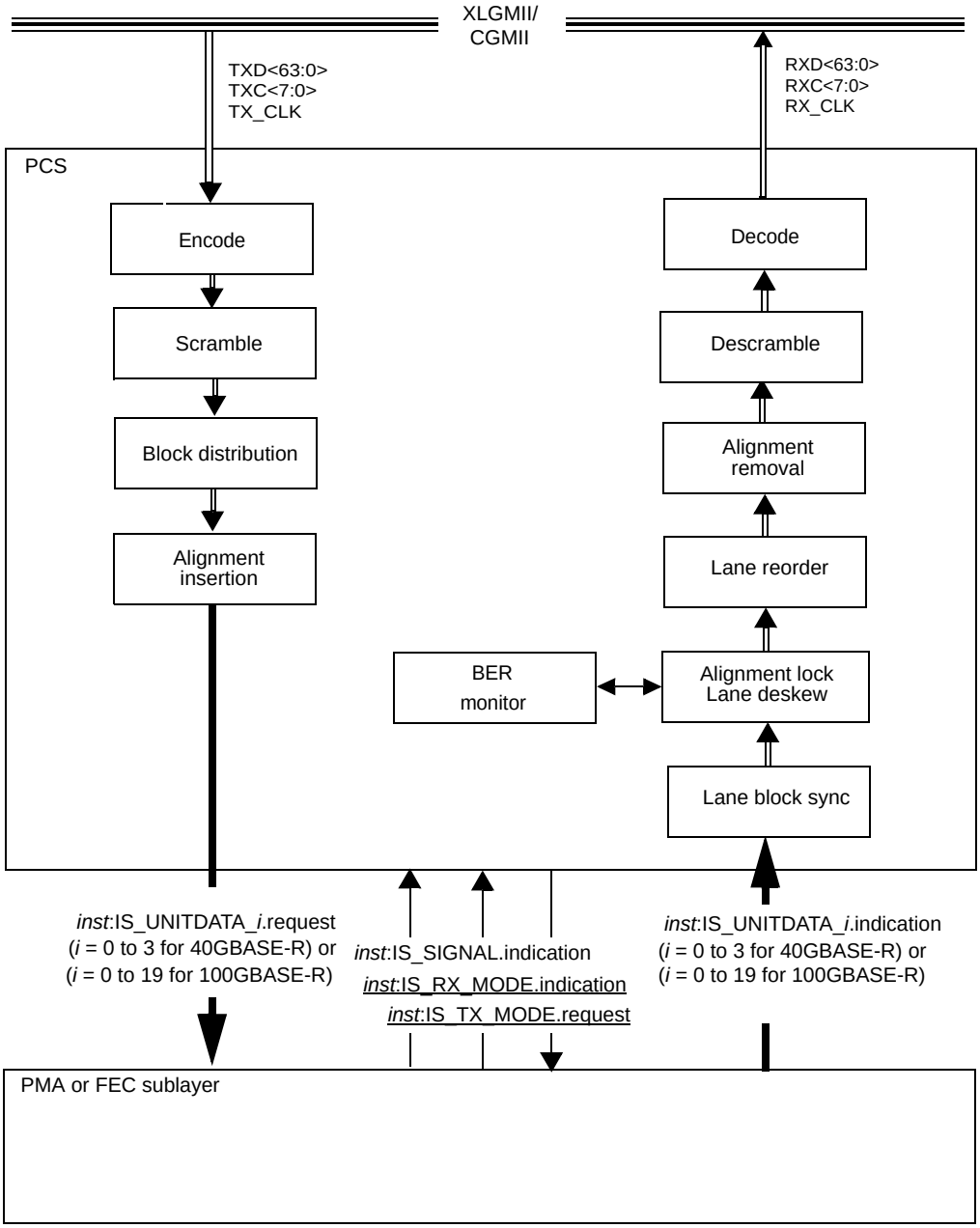


Figure 82–2—Functional block diagram

/I/ insertion and deletion shall occur in groups of 8. /I/s may be added following idle control characters or ordered sets. They shall not be added while data is being received.

To communicate LPI, the LPI control character /LI/ is sent continuously in place of /I/. LPI control characters are transmitted when LPI control characters are received from the CGMII. LPI characters may be added or deleted by the PCS to adapt between clock rates in a similar manner to idle control characters. /LI/ insertion and deletion shall occur in groups of 8. /LI/s may only be inserted following other LPI characters.

Insert 82.2.8a after 82.2.8 for RAM definition:

82.2.8a Rapid alignment marker insertion

For the optional EEE function, an alternate method of alignment is used. Rapid Alignment Markers (RAMs) function in a similar manner to the alignment markers described in 82.2.7. RAMs are sent in the place of normal alignment markers when the transmitter has an LPI transmit state other than TX_ACTIVE and down_count_done = FALSE. Additionally, the BIP component defined for alignment markers is replaced by a count down field (CD) so that the transition from RAMs to normal alignment markers can be indicated. The RAMs shall be inserted after every 7 66-bit blocks on each PCS lane. RAM insertion is performed in the same manner as shown in Figure 82–7 and Figure 82–9a. The transition from RAMs to normal alignment markers is shown in Figure 82–9a. The count down field is also used to communicate some of the states of the tx_mode when it is not being used to coordinate the transition.

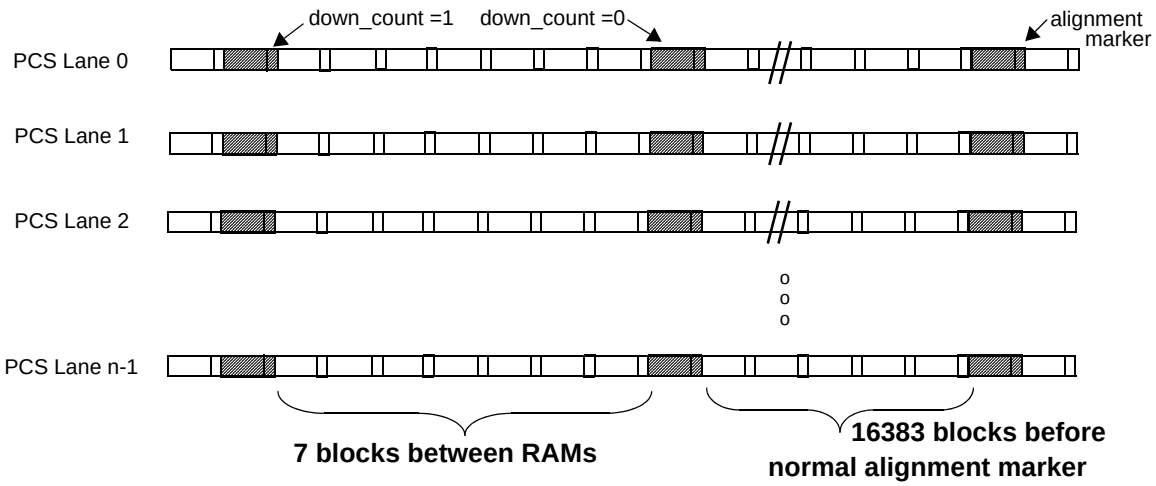


Figure 82–9a—RAM transition

The format of the RAMs is shown in Figure 82–9b.

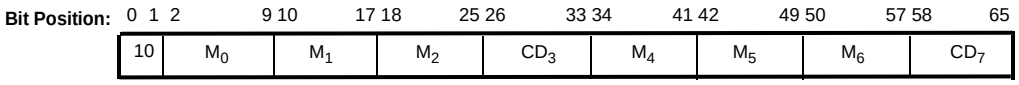


Figure 82–9b—RAM format

The content of the RAMs shall be as shown in Table 82–2 with CD_3 replacing BIP_3 and CD_7 replacing BIP_7. The contents depend on the PCS lane number and the octet number. Note that M_4 through M_6 are the bit-wise inversion of M_0 through M_2, respectively. Also CD_7 is the bit-wise inversion of CD_3. This property allows the alignment markers to be DC balanced. Lane markers 0 to 19 from Table 82–2 are used for the 100GBASE-R PCS. As an example, the lane marker for 100GBASE-R lane number 0 is sent as (left most bit sent first):

10 10000011 00010110 10000100 CD_3 01111100 11101001 01111011 CD_7

After the RAMs are inserted, data is sent to the PMA or FEC sublayer adjacent to the PCS.

The value of the CD_3 field is derived by the bit-wise XOR of the down_count variable with the M_0 value for the lane (therefore the last 5 RAMs on PCS lane 0 would have CD_3 values: 0xC5, 0xC2, 0xC3, 0xC0, 0xC1; for PCS lane 1 these would be: 0x99, 0x9E, 0x9F, 0x9C, 0x9D). The CD_7 field is the bit-wise inversion of CD_3. The CD field is used by the link partner to understand the expected transition from RAMs to normal AMs. It may also be used by a detached transmit PMA sublayer to infer the state of the PCS.

BIP statistics are only updated when the receiver is in the DATA state. In all other states, the running parity is not calculated.

Change 82.2.11 and Figure 82-10 for LPI override of synchronization:

82.2.11 Block synchronization

When the receive channel is operating in normal mode, the block synchronization function receives data via 4 (for 40GBASE-R) or 20 (for 100GBASE-R) IS_UNITDATA_*i*.indication primitives. The PCS forms 4 or 20 bit streams from the primitives by concatenating the bits from the indications of each primitive in order from each *inst*:IS_UNITDATA_0.indication to *inst*:IS_UNITDATA_3.indication or *inst*:IS_UNITDATA_0.indication to *inst*:IS_UNITDATA_19.indication. It obtains lock to the 66-bit blocks in each bit stream using the sync headers and outputs 66-bit blocks. Block lock is obtained as specified in the block lock state diagram shown in Figure 82–10.

If EEE is not supported then block_lock is identical to rx_block_lock. Otherwise the relationship between block_lock and rx_block_lock is given by Figure 82–17 the LPI receive state diagram.

82.2.12 PCS lane deskew

Insert the following at the end of 82.2.12, change Figures 82-11 and 82-12 for LPI override of align status:

If EEE is not supported then align_status is identical to rx_align_status. Otherwise the relationship between align_status and rx_align_status is given by Figure 82–17 the LPI receive state diagram.

82.2.18.2.2 Variables

Insert a note in 82.2.18.2.2 below the definition for "align_status":

NOTE: If the EEE capability is supported, then this variable is affected by the LPI receive state diagram. If the EEE capability is not supported then this variable is identical to rx_align_status controlled by the Alignment marker lock state diagram.

Insert a note in 82.2.18.2.2 below the definition for "block_lock":

NOTE: If the EEE capability is supported, then this variable is affected by the LPI receive state diagram. If the EEE capability is not supported then this variable is identical to rx_block_lock controlled by the Block lock state diagram.

Insert the following new variables at appropriate places in 82.2.18.2.2:

rx_align_status
: Variable used by the PCS lane deskew process to reflect the status of the PCS lane-to-lane alignment. This variable is set true when all lanes are synchronized and aligned, set false when the deskew process is not complete.

rx_block_lock
Variable used by the Block lock state diagram to reflect the status of the code-group delineation. This variable is set true when the receiver acquires block delineation.

Insert the following new text and variables at the end of the existing subclause 82.2.18.2.2:

The following variables are used only for the EEE capability.

down_count_enable
Boolean variable controlling decrement of the counter down_count. This variable is set by the LPI transmit state diagram.

LPI_FW
Boolean variable controlling the wake mode for the LPI transmit function. This variable is set true when the transmitter is to use the Fast Wake mechanism, and false otherwise.

received_tx_mode
Variable reflecting state of the LPI transmit function for the link partner. The value of this variable is inferred from the coding of the RAMs of the incoming data stream.

rx_lpi_active
A Boolean variable that is set to true when the receiver is in a low power state and set to false when it is in an active state and capable of receiving data.

Rx LPI indication:
A boolean variable indicating the current state of the receive LPI function. This flag is set to true (register bit set to one) when the LPI receive state diagram is in any state other than RX_ACTIVE. This status is reflected in MDIO register 3.1.8. A latch high view of this status is reflected in MDIO register 3.1.10 (Rx LPI received).

rx_mode
A parameter generated by the PMA/PMD sublayer to reflect the state of the received signal.The parameter may have one of four values : DATA, QUIET, or ALERT.

Tx LPI indication:
A boolean variable indicating the current state of the transmit LPI function. This flag is set to true (register bit set to one) when the LPI transmit state diagram is in any state other than TX_ACTIVE. This status is reflected in MDIO register 3.1.9. A latch high view of this status is reflected in MDIO register 3.1.11 (Tx LPI received).

tx_mode
A variable reflecting state of the LPI transmit function as described by the LPI transmit state diagram (Figure 82–16). When tx_mode is set to QUIET sublayer may go into a low power state.

82.2.18.2.3 Functions

Change 89.2.18.2.3 function definitions for LPI block types:

AM_SLIP
Causes the next candidate block position to be tested. The precise method for determining the next candidate block position is not specified and is implementation dependent. However, an implementation shall ensure that all possible blocks are evaluated.

DECODE(rx_coded<65:0>)

Decodes the 66-bit vector returning rx_raw<71:0>, which is sent to the XLGMII/CGMII. The DECODE function shall decode the block as specified in 82.2.3.

ENCODE(tx_raw<71:0>)

Encodes the 72-bit vector returning tx_coded<65:0> of which tx_coded<65:2> is sent to the scrambler. The two bits of the sync header bypass the scrambler. The ENCODE function shall encode the block as specified in 82.2.3.

R_TYPE(rx_coded<65:0>)

This function classifies the current rx_coded<65:0> vector as belonging to one of the following ~~five~~ types, depending on its contents. The classification results are returned via the r_block_type variable.

Values: C; The vector contains a sync header of 10 and one of the following:
a) A block type field of 0x1E and eight valid control characters other than /E/ or /LI/;
b) A block type field of 0x4B.
LI; For EEE capability, the LI type is supported where the vector contains a sync header of 10, a block type field of 0x1e and eight control characters of 0x06 (/LI/).
S; The vector contains a sync header of 10 and the following:
a) A block type field of 0x78.
T; The vector contains a sync header of 10, a block type field of 0x87, 0x99, 0xAA, 0xB4, 0xCC, 0xD2, 0xE1 or 0xFF and all control characters are valid.
D; The vector contains a sync header of 01.
E; The vector does not meet the criteria for any other value.

Valid control characters are specified in Table 82–1.
Note: A PCS that does not support EEE, classifies vectors containing one or more /LI/ control characters as type E.

R_TYPE_NEXT

This function classifies the 66-bit rx_coded vector that immediately follows the current rx_coded<65:0> vector as belonging to one of the five types defined in R_TYPE, depending on its contents. It is intended to perform a prescient end of packet check. The classification results are returned via the r_block_type_next variable.

SLIP

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

T_TYPE = (tx_raw<71:0>)

This function classifies each 72-bit tx_raw vector as belonging to one of the following ~~five~~ types depending on its contents. The classification results are returned via the t_block_type variable.

Values: C; The vector contains one of the following:
a) Eight valid control characters other than /O/, /S/, /T/ and /E/ or /LI/;
b) One valid ordered set.
LI; For EEE capability, this vector contains eight /LI/ characters.
S; The vector contains an /S/ in its first character, and all characters following the /S/ are data characters.
T; The vector contains a /T/ in one of its characters, all characters before the /T/ are data characters, and all characters following the /T/ are valid control characters other than /O/, /S/ and /T/.
D; The vector contains eight data characters.
E; The vector does not meet the criteria for any other value.

A tx_raw character is a control character if its associated TXC bit is asserted. A valid control character is one containing an XLGMII/CGMII control code specified in Table 82–1. A valid ordered_set consists of a valid /O/ character in the first character and data characters in the seven characters following the /O/. A valid /O/ is any character with a value for O code in Table 82–1.
Note: A PCS that does not support EEE, classifies vectors containing one or more /LI/ control characters as type E.

82.2.18.2.4 Counters

Change the definition for am_counter in 82.2.18.2.4 for RAMs:

am_counter

This counter counts 16383 66-bit blocks that separate two consecutive alignment markers for normal alignment markers or 7 66-bit blocks for rapid alignment markers for the optional EEE capability.

Insert new counters into 82.2.18.2.4, and new timers into 82.2.18.2.5 in support of the LPI state diagrams. In each case, insert the new text at the end of the existing subclause:

The following counters are used only for the EEE capability.

down_count

A counter that is used in rapid alignment markers and is decremented each time a RAM is sent while variable down_count_enable = true. The counter initial value is set by the LPI transmit state diagram. When the down_count counter reaches zero it will set the variable down_count_done = true.

wake_error_counter

A counter that is incremented each time that the LPI receive state diagram enters the RX_WTF state indicating that a wake time fault has been detected. The counter is reflected in register 3.22 (see 45.2.3.10)

82.2.18.2.5 Timers

The following timers are used only for the EEE capability.

one_us_timer

A timer used to count approximately 1 μs intervals. The timer terminal count is set to T_{1U}. When the timer reaches terminal count it will set the one_us_timer_done = true.

rx_tq_timer

This timer is started when the PCS receiver enters the RX_SLEEP state. The timer terminal count is set to T_{QR}. When the timer reaches terminal count it will set the rx_tq_timer_done = true.

rx_tw_timer

This timer is started when the PCS receiver enters the RX_WAKE state. The timer terminal count shall be set to a value no larger than the maximum value given for T_{WR} in TABLE 82–5. When the timer reaches terminal count it will set the rx_tw_timer_done = true.

rx_wf_timer

This timer is started when the PCS receiver enters the RX_WTF state, indicating that the receiver has encountered a wake time fault. The rx_wf_timer allows the receiver an additional period in which to synchronize or return to the QUIET state before a link failure is indicated. The timer terminal count is set to T_{WTF}. When the timer reaches terminal count it will set the rx_wf_timer_done = true.

tx_ts_timer

This timer is started when the PCS transmitter enters the TX_SLEEP state. The timer terminal count is set to T_{SL}. When the timer reaches terminal count it will set the tx_ts_timer_done = true.

tx_tq_timer

This timer is started when the PCS transmitter enters the TX_QUIET state. The timer terminal count is set to T_{QL}. When the timer reaches terminal count it will set the tx_tq_timer_done = true.

tx_tw_timer

This timer is started when the PCS transmitter enters the TX_WAKE or TX_RF_WAKE state. The timer terminal count is set to T_{WL}. When the timer reaches terminal count it will set the tx_tw_timer_done = true.

82.2.18.3 State diagrams

Insert 82.2.18.3.1 at the end of 82.2.18.3:

82.2.18.3.1 LPI state diagrams

A PCS which supports the EEE capability shall implement the LPI transmit and receive processes as shown in figures 82–16 and 82–17. The transmit LPI state diagram controls tx_mode which disables the transmitter when it is set to QUIET. The receive LPI state diagram controls block_lock during LPI and signals the end of LPI to the receive state diagram.

Following a period of LPI, the receiver is required to achieve block synchronization within the wakeup time specified (See Figure 82–17). The implementation of the block synchronization state diagram should use techniques to ensure that block lock is achieved with minimal numbers of slip attempts. If Fast Wake is selected then the receiver is expected to maintain sufficient state to allow much faster wake up.

The LPI functions shall use timer values for these state diagrams as shown in TABLE 82–5a for transmit and TABLE 82–5b for receive.

Table 82–5a—Transmitter LPI timing parameters

Parameter	Description	Min	Max	Units
T_{SL}	Local Sleep Time from entering the TX_SLEEP state to when tx_mode is set to QUIET, LPI_FW = FALSE	0.9	1.1	µs
T_{SL}	Local Sleep Time from entering the TX_SLEEP state to when tx_mode is set to FW, LPI_FW = TRUE	240	260	ns
T_{QL}	Local Quiet Time from when tx_mode is set to QUIET or FW to entry into the TX_RF_ALERT or TX_WAKE states	1.7	1.8	ms
T_{WL}	Time spent in the TX_WAKE or TX_RF_WAKE states, LPI_FW = FALSE	3.9	4.1	µs
T_{WL}	Time spent in the TX_WAKE state, LPI_FW = TRUE	240	260	ns
T_{1U}	Time spent in the TX_ALERT or TX_RF_ALERT states	1.1	1.3	µs

Table 82–5b—Receiver LPI timing parameters

Parameter	Description	Min	Max	Units
T_{QR}	The time the receiver waits for rx_mode to be set to ALERT or DATA while in the RX_SLEEP and RX_QUIET or RX_FW states before asserting receive fault	2	3	ms
T_{WR}	Time the receiver waits in the RX_WAKE state before indicating a wake time fault, LPI_FW = FALSE.	—	4.5	µs
T_{WR}	Time the receiver waits in the RX_WAKE state before indicating a wake time fault, LPI_FW = TRUE.	—	300	ns
T_{WTF}	Wake time fault recovery time	—	10	ms

82.3.1 PMD MDIO function mapping

Add the following row to Table 82-6 in 82.3.1:

Table 82–6—MDIO/PMD control variable mapping

MDIO control variable	PCS register name	Register/ bit number	PCS control variable
LPI_FW	LPI fast wake enable	TBD	LPI_FW

Add the following rows to Table 82-7 in 82.3.1:

Table 82–7—MDIO/PMD status variable mapping

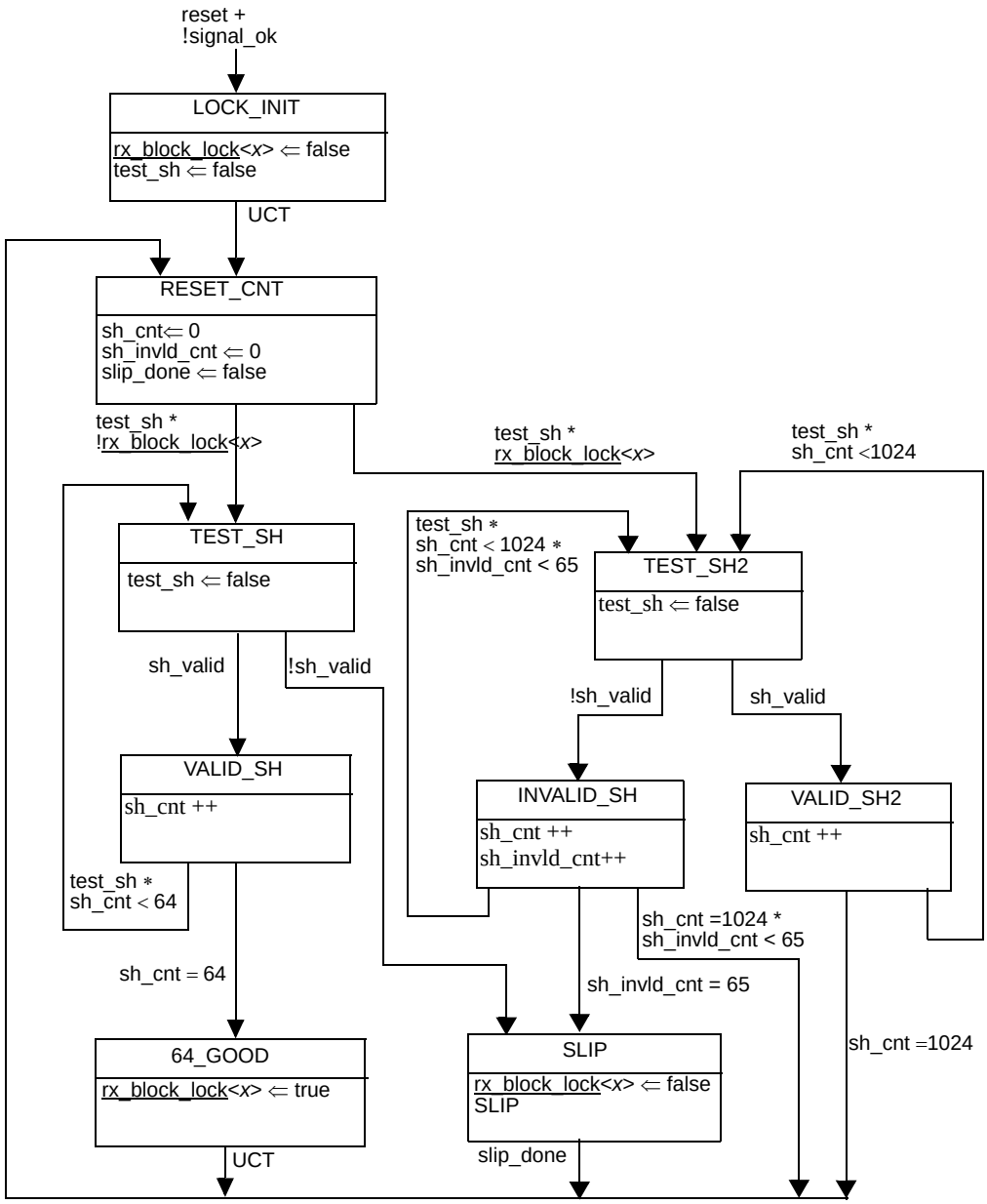
MDIO status variable	PCS register name	Register/ bit number	PCS status variable
Tx LPI indication	Tx LPI indication	3.1.9	Tx LPI indication
Tx LPI received	Tx LPI received	3.1.11	Tx LPI received
Rx LPI indication	Rx LPI indication	3.1.8	Rx LPI indication
Rx LPI received	Rx LPI received	3.1.10	Rx LPI received
Wake_error_counter	Wake_error_counter	3.22	Wake_error_counter

82.6 Auto-Negotiation

Change 82.6 to add new PHY types (per marris_01_0312.pdf)

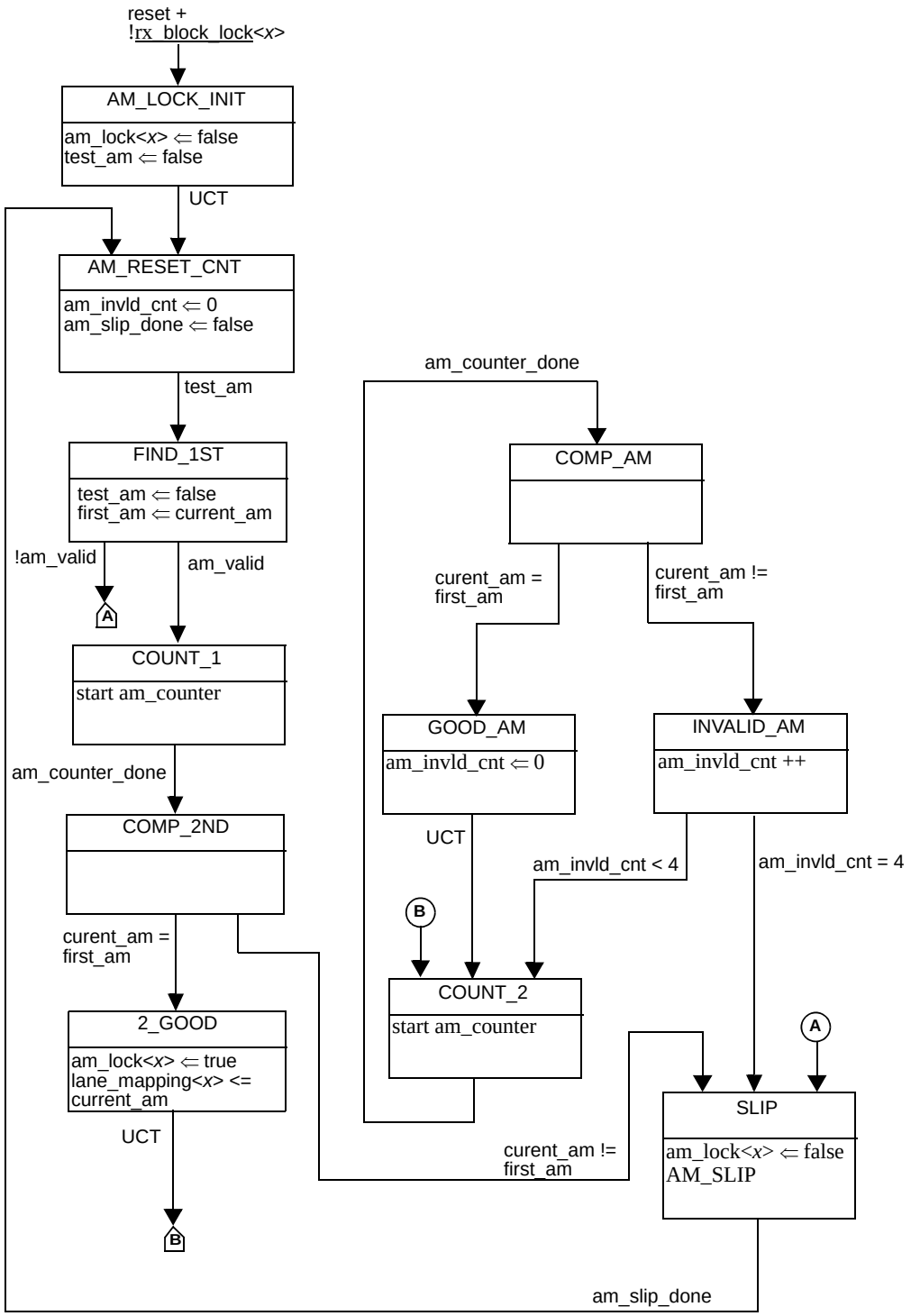
The following requirements apply to a PCS used with a 40GBASE-KR4 PMD, 40GBASE-CR4 PMD, ~~or~~ 100GBASE-CR10, 100GBASE-CR4, 100GBASE-KR4, or 100GBASE-KP4 PMD where support for the Auto-Negotiation process defined in Clause 73 is mandatory. The PCS shall support the primitive AN_LINK.indication(link_status) (see 73.9). The parameter link_status shall take the value FAIL when

PCS_status=false and the value OK when PCS_status=true. The primitive shall be generated when the value of link_status changes.



NOTE— rx_block_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)

Figure 82–10—Block lock state diagram



NOTE— 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)

Figure 82–11—Alignment marker lock state diagram

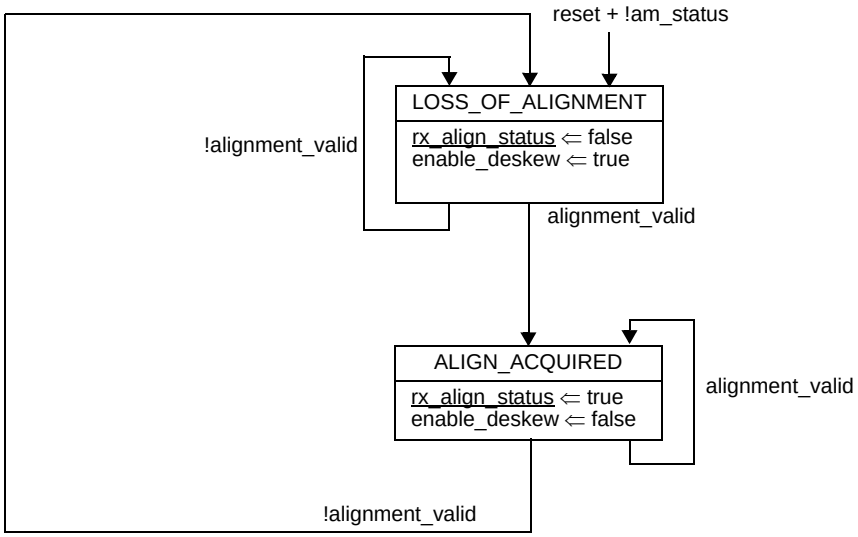


Figure 82–12—PCS deskew state diagram

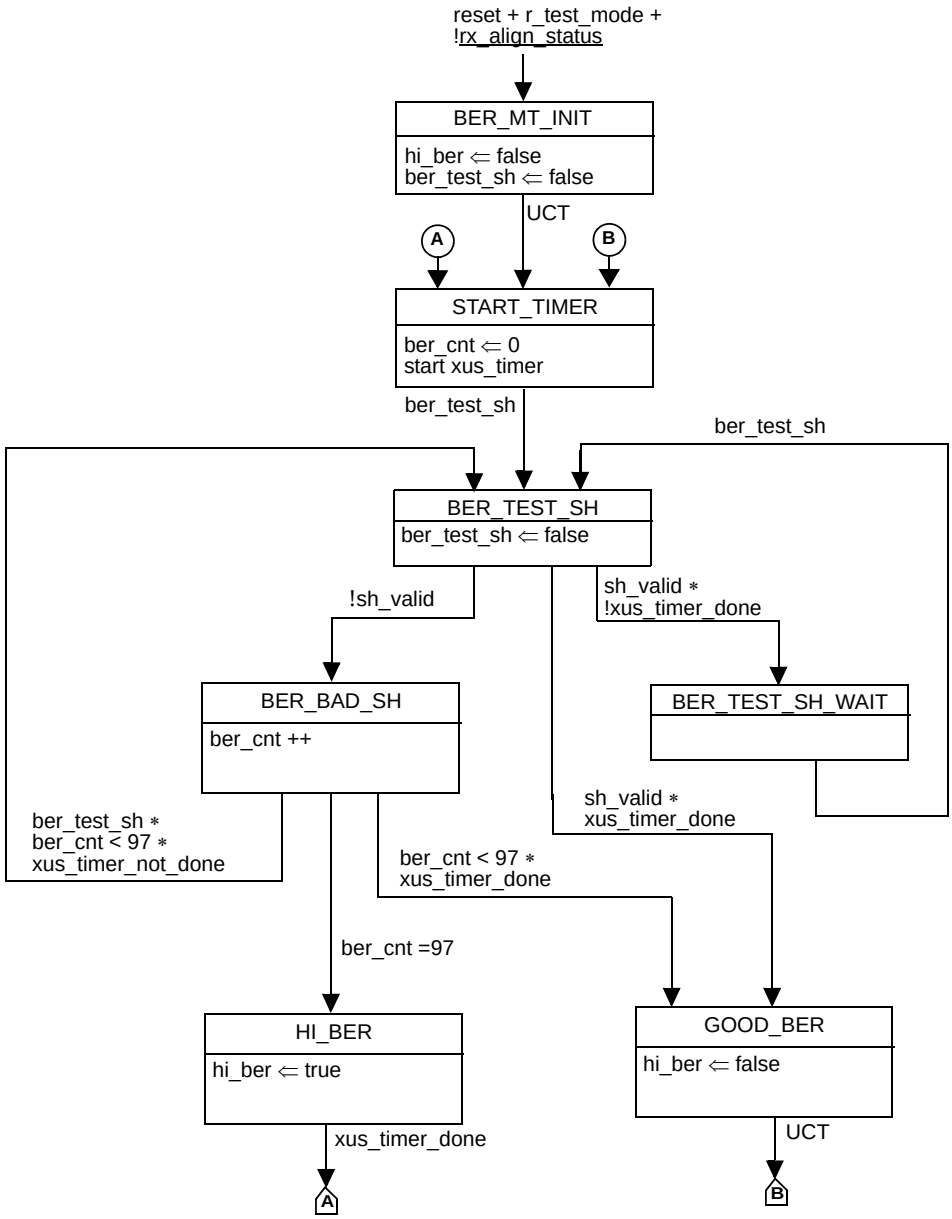


Figure 82–13—BER monitor state diagram

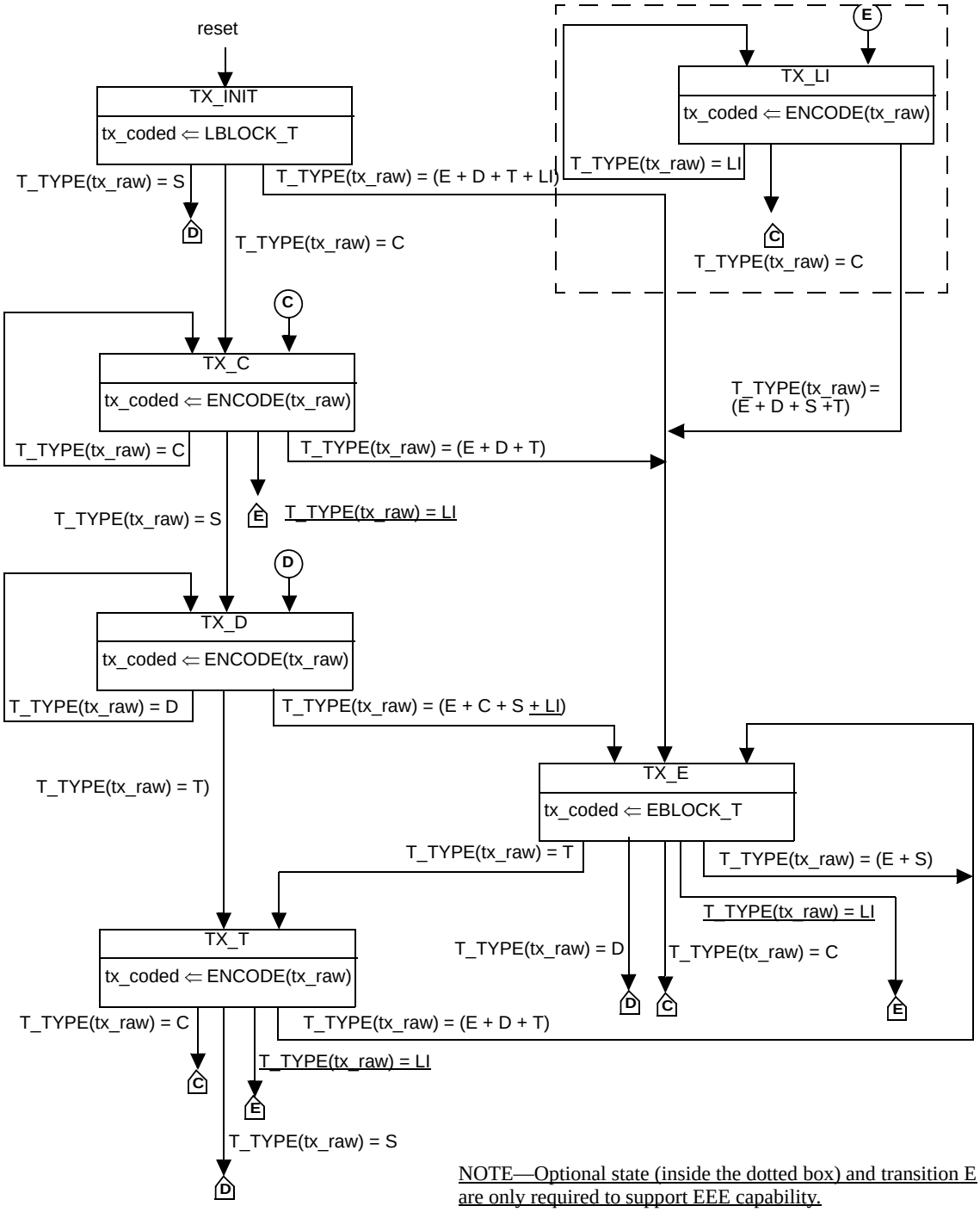
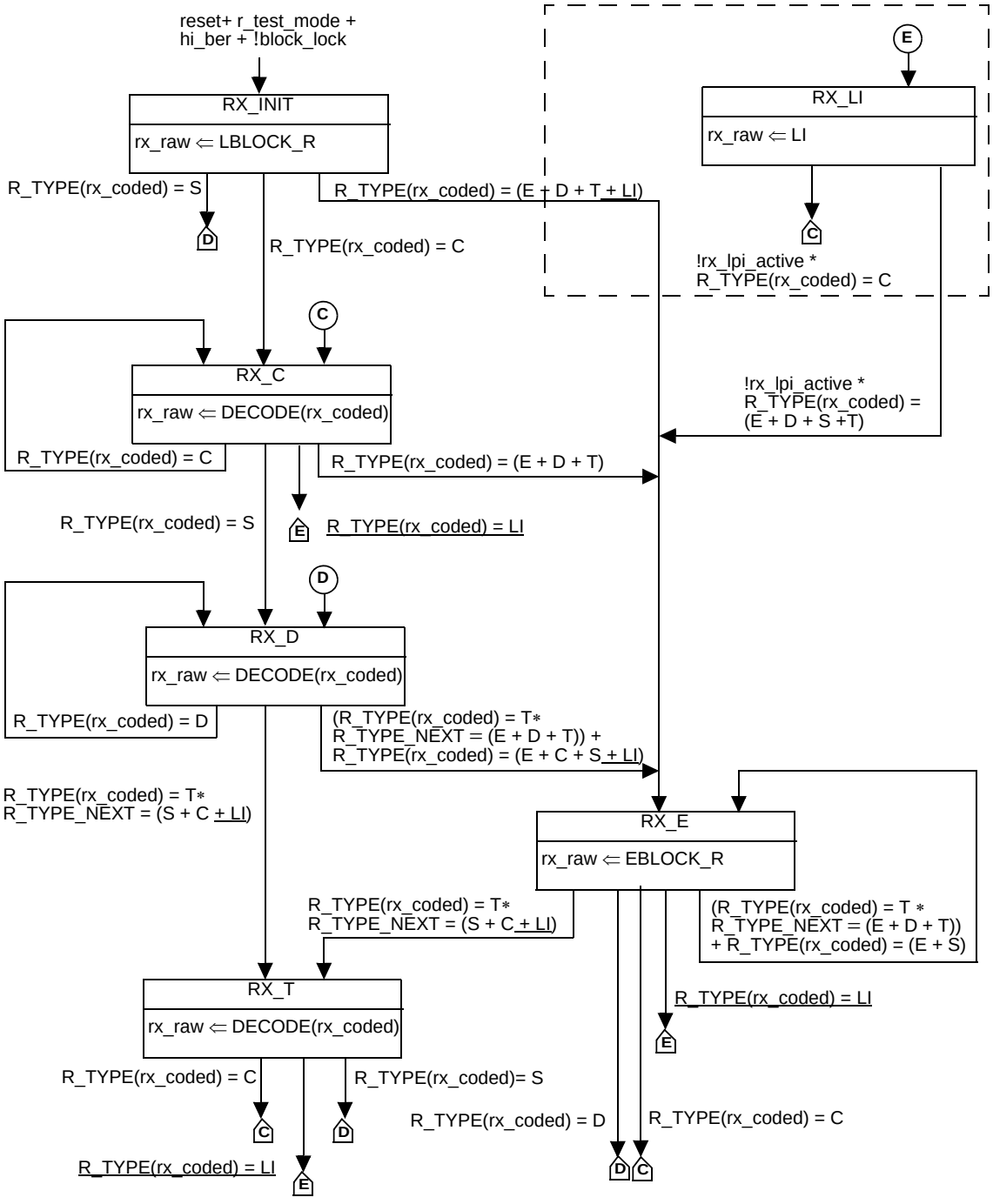


Figure 82–14—Transmit state diagram



NOTE—Optional state (inside the dotted box) and transition E are only required to support EEE capability.

Figure 82–15—Receive state diagram

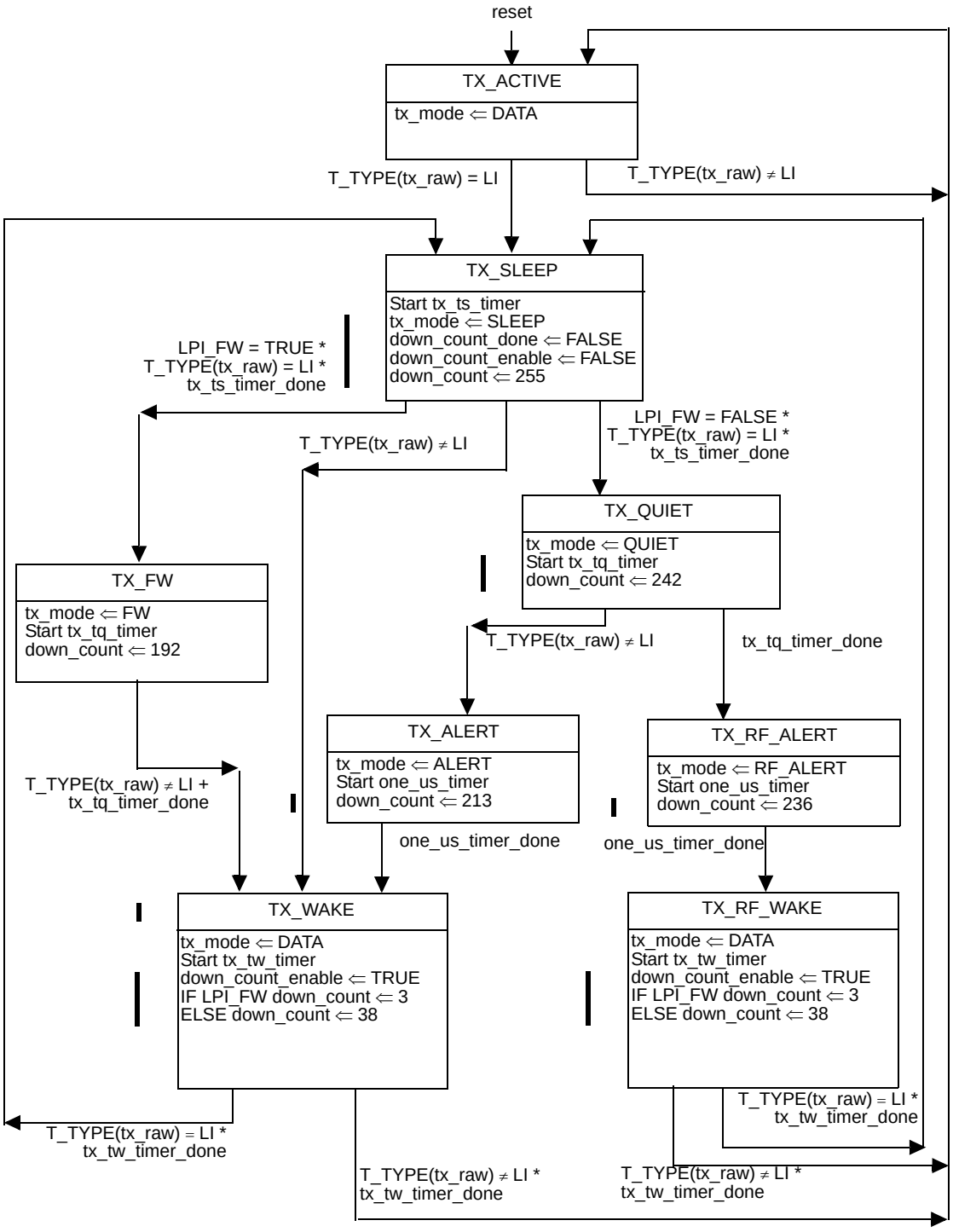


Figure 82–16—LPI Transmit state diagram

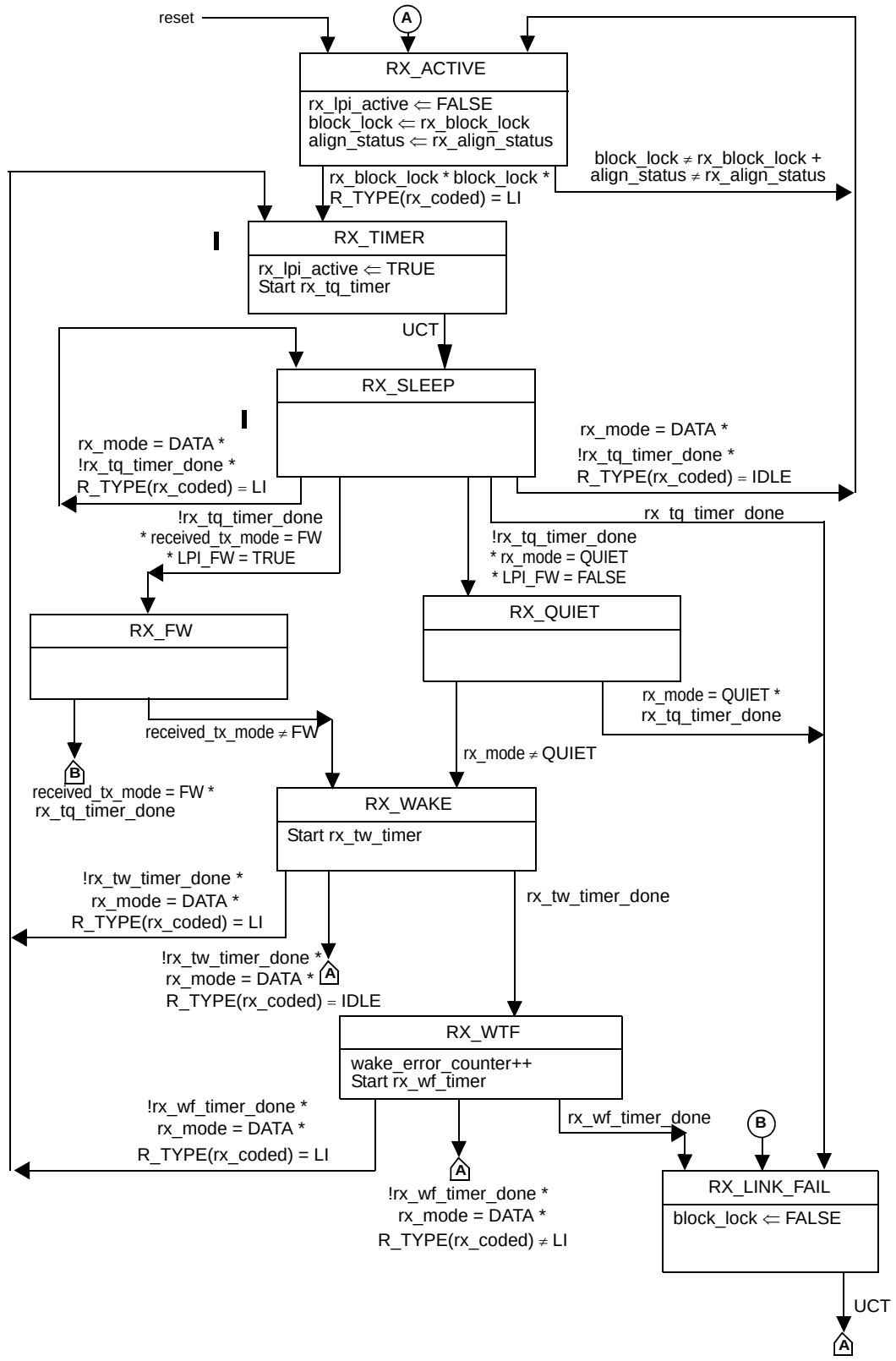


Figure 82–17—LPI Receive state diagram

82.7 Protocol implementation conformance statement (PICS) proforma for Clause 82, Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R[6]

Insert the following row into table 82.7.3:

82.7.3 Major Capabilities/Options

Item	Feature	Subclause	Value/Comment	Status	Support
*LPI	Implementation of LPI	82.2.3.4		O	Yes [ ] No [ ]

Change rows AN1* and AN2 in 82.7.6.5:

82.7.6.5 Auto-Negotiation for Backplane Ethernet functions

Item	Feature	Subclause	Value/Comment	Status	Support
AN1*	Support for use with a 40GBASE-KR4, 40GBASE-CR4, ~~or~~ 100GBASE-CR10 PMD, 100GBASE-CR4, 100GBASE-KR4, or 100GBASE-KP4	82.6	AN technology dependent interface described in Clause 73	O	Yes [ ]
AN2	AN_LINK.indication primitive	82.6	Support of the primitive AN_LINK.indica-tion(link_status), when the PCS is used with ~~10GBASE-KR PMD~~ 40GBASE-KR4, 40GBASE-CR4, 100GBASE-CR10 PMD, 100GBASE-CR4, 100GBASE-KR4, or 100GBASE-KP4	AN1:M	Yes [ ]

Insert the new subclause 82.7.6.6 after 82.7.6.5 for LPI functions:

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

82.7.6.6 LPI functions

Item	Feature	Subclause	Value/Comment	Status	Support
LP-01	Insertion and deletion of LPIs in groups of 8	82.2.3.6		LPI:M	Yes [ ] No [ ]
LP-05	Transmit state diagrams	82.2.18.3	Support LPI operation in Figure 82–14	LPI:M	Yes [ ] No [ ]
LP-06	Receive state diagrams	82.2.18.3	Support LPI operation in Figure 82–15	LPI:M	Yes [ ] No [ ]
LP-07	LPI transmit state diagrams	82.2.18.3.1	Meets the requirements of Figure 82–16	LPI:M	Yes [ ] No [ ]
LP-08	LPI receive state diagrams	82.2.18.3.1	Meets the requirements of Figure 82–17	LPI:M	Yes [ ] No [ ]
LP-09	LPI transmit timing	82.2.18.3.1	Meets the requirements of Table 82–5a	LPI:M	Yes [ ] No [ ]
LP-10	LPI receive timing	82.2.18.3.1	Meets the requirements of Table 82–5b	LPI:M	Yes [ ] No [ ]

83. Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

83.1 Overview

Change NOTE 1 in Figure 83-1 as shown:

NOTE1—~~OPTIONAL OR OMITTED DEPENDING~~ CONDITIONAL BASED ON PHY TYPE

Change the first paragraph of 83.3 as follows:

83.1.1 Scope

This clause specifies the Physical Medium Attachment sublayer (PMA) that is common to two families of (40 Gb/s and 100 Gb/s) Physical Layer implementations, known as 40GBASE-R and 100GBASE-R. The PMA allows the PCS (specified in Clause 82) to connect in a media-independent way with a range of physical media. The 40GBASE-R PMA(s) can support any of the 40 Gb/s PMDs in Table 80–2, except 100GBASE-KP4 (Clause 94). The 100GBASE-R PMA(s) can support any of the 100 Gb/s PMDs in ~~Table 80–2~~Table 80–2a. The terms 40GBASE-R and 100GBASE-R are used when referring generally to Physical Layers using the PMA defined in this clause.

Add the following at the end of 83.3 for the EEE service interface:

83.3 PMA service interface

If the optional Energy Efficient Ethernet (EEE) capability is supported (see Clause 78, 78.3) then the inter-sublayer service interface includes two additional primitives defined as follows:

IS_TX_MODE.request
IS_RX_MODE.indication

The IS_TX_MODE.request primitive is used to communicate the state of the PCS LPI transmit function to other sublayers in the PHY. The IS_RX_MODE.indication primitive is used to communicate the state of the received signal to the PCS (and other sublayers of the PHY).

A physically instantiated service interface with the optional Energy Efficient Ethernet (EEE) capability (see 78.3) may enter a low power state to conserve energy during periods of low link utilization. The ability to support transition to a low power state in the ingress direction is indicated by register TBD (PMA Ingress AUI Stop Ability, PIASA) and register TBD for the egress direction (PMA Egress AUI Stop Ability,

PEASA). Transition to the low power state is enabled in the ingress direction by register TBD (PMA Ingress AUI Stop Enable, PIASE) and register TBD for the egress direction (PMA Egress AUI Stop Enable, PEASE). The system shall not assert the enable bit for an interface unless the corresponding ability bit at the other side of the interface is also asserted. If the PIASE bit is TRUE, then the PMA may disable transmitters on the physical instantiation of the ingress AUI when rx_mode is QUIET. If the PEASE bit is TRUE, then the PMA may disable transmitters on the physical instantiation of the egress AUI when tx_mode is QUIET.

83.6 PMA MDIO function mapping

Add rows to Table 83-2 for fast wake and stop enable:

Table 83–2—MDIO/PMA control variable mapping

MDIO variable	PMA/PMD register name	Register/ bit number	PMA control variable
LPI_FW	LPI fast wake enable	TBD	LPI_FW
PIASE	PMA ingress AUI stop enable	TBD	PIASE
PEASE	PMA egress AUI stop enable	TBD	PEASE

Add rows to Table 83-3 for stop ability:

Table 83–3—MDIO/PMA status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMA status variable
PIASA	PMA ingress AUI stop ability	TBD	PIASA
PEASA	PMA egress AUI stop ability	TBD	PEASA

83.7 Protocol implementation conformance statement (PICS) proforma for Clause 83, Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R[7]

Insert the following row into table 82.7.3:

83.7.3 Major Capabilities/Options

Item	Feature	Subclause	Value/Comment	Status	Support
*LPI	Implementation of LPI	83.3		O	Yes [ ] No [ ]

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

84. Physical Medium Dependent sublayer and baseband medium, type 40GBASE-KR4

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

Proposals for amendments to this clause that define operation with the optional Energy Efficient Ethernet capability are needed.

85. Physical Medium Dependent sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

85.1 Overview

Add a row in Table 85-1 for EEE:

Table 85–1—Physical Layer clauses associated with the 40GBASE-CR4 and 100GBASE-CR10 PMDs

Associated clause	40GBASE-CR4	100GBASE-CR10
78—Energy Efficient Ethernet	Not applicable	Optional

Insert the following at the end of 85.1:

A 100GBASE-CR10 PHY with the optional Energy Efficient Ethernet (EEE) capability may optionally enter the Low Power Idle (LPI) mode to conserve energy during periods of low link utilization (see Clause 78).

85.2 Physical Medium Dependent (PMD) service interface

Add the following at the end of 85.2 for the EEE service interface

If the optional Energy Efficient Ethernet (EEE) capability is supported (see Clause 78, 78.3) then the inter-sublayer service interface includes two additional primitives defined as follows:

PMD:IS_TX_MODE.request
PMD:IS_RX_MODE.indication

The TX_MODE parameter takes on one of eight values: DATA, SLEEP, QUIET, FW, ALERT, RF_ALERT, WAKE or RF_WAKE. When TX_MODE = QUIET, transmission is disabled; when TX_MODE = ALERT or RF_ALERT, the alert signal is transmitted.

The RX_MODE parameter takes on one of four values: DATA, QUIET, or ALERT. This parameter reflects the status of the received signal.

85.7.2 PMD Transmit function

Add the following at the end of 85.7.2 for the EEE function:

If the optional Energy Efficient Ethernet (EEE) capability is supported (see Clause 78) then when tx_mode is set to ALERT, the PMD will transmit a repeating 16-bit pattern, hexadecimal 0xFF00. When tx_mode is ALERT, the transmitter equalizer taps are set to the preset state specified in 85.8.3.3.1. When tx_mode is QUIET, the transmitter is disabled as specified in 85.7.6. For all other states of tx_mode, the driver coefficients are restored to their states resolved during training.

85.7.4 Global PMD signal detect function

Editor's note (to be removed prior to final publication):

The following behavior has been proposed but not adopted by the Task Force for EEE behavior:

When the PHY supports the optional EEE capability, the signal detect function is also used to control the state of the rx_mode parameter. The parameter rx_mode is set to DATA following system reset or completion of training. Following the reception of a data stream containing RAMs with the code indicating tx_mode = SLEEP, rx_mode shall be set to QUIET and shal remain in that state until a signal is detected at the receiver input that is the output of a channel that satisfies the requirements of all the parameters of both interference tolerance test channels defined in 72.7.2.1 when driven by a square wave pattern with a period of 16 unit intervals and peak-to-peak differential output amplitude of 720 mV. Parameter rx_mode shall be set to ALERT within 500ns of the application of this signal. Parameter rx_mode shall return to DATA within 4uS of a return to normal data reception.

85.7.6 Global PMD transmit disable function

Change 85.7.6 for the EEE transmit_disable:

The Global_PMD_transmit_disable function is mandatory if EEE is supported and is otherwise optional. When implemented, it allows all of the transmitters to be disabled with a single variable.

a) When Global_PMD_transmit_disable variable is set to one, this function shall turn off all of the transmitters such that each transmitter drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage in Table 85–2.
b) If a PMD_fault (85.7.9) is detected, then the PMD may turn off the electrical transmitter in all lanes.
c) Loopback, as defined in 85.7.8, shall not be affected by Global_PMD_transmit_disable.
d) For EEE capability, the PMD_transmit_disable function shall turn off the transmitter after tx_mode is set to QUIET within a time and voltage level specified in 72.7.1.4. The PMD_transmit_disable function shall turn on the transmitter after tx_mode is set to DATA or ALERT within the time and voltage level specified in 72.7.1.4.

85.8.3 Transmitter characteristics

Add the following row in Table 85-5 in 85.8.3:

Table 85–2—Transmitter characteristics at TP2 summary

Parameter	Subclause reference	Value	Units
Common-mode voltage deviation (max) during LPI	72.7.1.4	150	mV

85.13 Protocol implementation conformance statement (PICS) proforma for Clause 85, Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10[8]

Insert the following row into table 85.13.3:

85.13.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*LPI	Implementation of LPI	85.2		O	Yes [ ] No [ ]

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

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

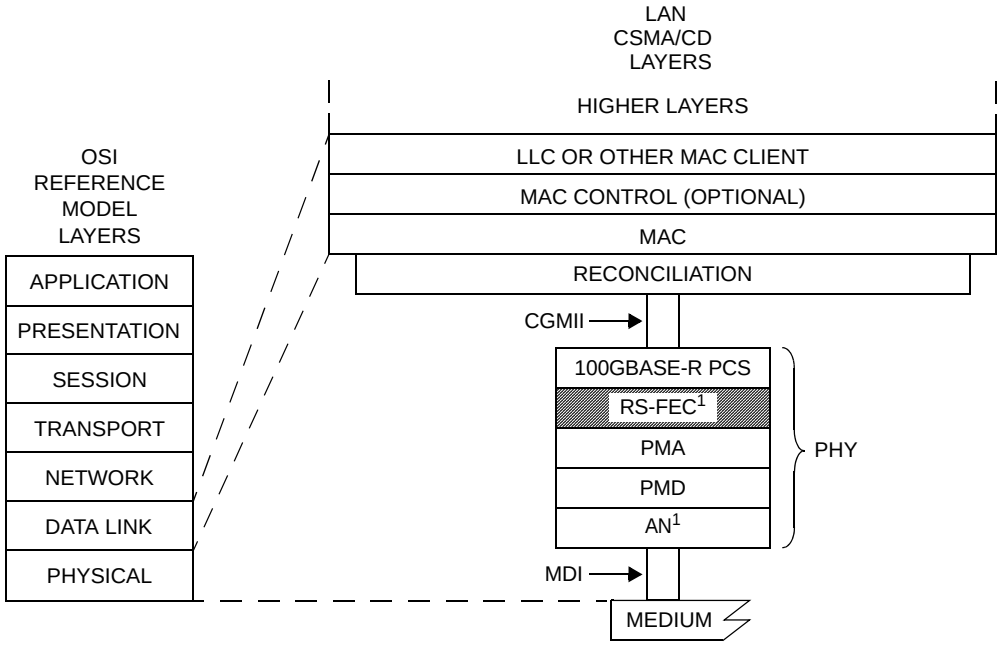
91.1 Overview

91.1.1 Scope

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

91.1.2 Position of RS-FEC in the 100GBASE-R sublayers

Figure 91–1 shows the relationship of the RS-FEC sublayer to the ISO/IEC Open System Interconnection (OSI) reference model.



AN = AUTO-NEGOTATION
CGMII = 100 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
RS-FEC = REED-SOLOMON FORWARD ERROR CORRECTION

NOTE 1—CONDITIONAL BASED ON PHY TYPE

Figure 91–1—RS-FEC relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

91.2 FEC service interface

This subclause specifies the services provided by the RS-FEC sublayer. The service interface is described in an abstract manner and does not imply any particular implementation.

The FEC service interface is provided to allow the PCS to transfer information to and from the RS-FEC. The PCS may be connected to the RS-FEC using an optional instantiation of the PMA service interface (refer to Annex 83A) in which case a PMA will be the client of the FEC service interface.

The FEC service interface is an instance of the inter-sublayer service interface defined in 80.3. The FEC service interface primitives are summarized as follows:

FEC:IS_UNITDATA_*i*.request
FEC:IS_UNITDATA_*i*.indication
FEC:IS_SIGNAL.indication

The RS-FEC operates on twenty parallel bit streams, hence i = 0 to 19. The PCS (or PMA) continuously sends twenty parallel bit streams to the RS-FEC, one per lane, each at a nominal signaling rate of 5.15625 GBd. The RS-FEC continuously sends twenty parallel bit streams to the PCS (or PMA), one per lane, each at a nominal signaling rate of 5.15625 GBd

This FEC:IS_SIGNAL.indication primitive is sent by the RS-FEC to the PCS (or PMA) to indicate the status of the RS-FEC receive function.

Editor's note (to be removed prior to final publication):

The adopted baseline proposal(s) do not define or imply how SIGNAL_OK is set to indicate the status of the RS-FEC receive function. One possible approach is to base SIGNAL_OK on the state of the synchronization state diagram. However, the baseline proposal(s) left this state diagram TBD so the assignment of SIGNAL_OK will also be left TBD until a more detailed proposal is adopted.

If the optional EEE capability is supported, then the FEC service interface includes two additional primitives as follows:

FEC:IS_TX_MODE.request
FEC:IS_RX_MODE.indicate

Editor's note (to be removed prior to final publication):

Per barrass_01_0312, the behavior of the RS-FEC for the optional EEE capability is TBD.

91.3 PMA compatibility

The RS-FEC sublayer requires that the PMA service interface consist of exactly four upstream lanes and exactly four downstream lanes. Therefore, the RS-FEC sublayer may be a client of the PMA sublayer defined in Clause 83 when the PMA service interface width, p, is set to 4. The RS-FEC sublayer may also be a client of the PMA sublayer defined in Clause 94.

91.4 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 TBD bit times (TBD pause_quanta or TBD ns). A description of overall sys-

tem delay constraints and the definitions for bit times and pause_quanta can be found in 80.4 and its references.

91.5 Functions within the RS-FEC sublayer

91.5.1 Functional block diagram

A functional block diagram of the RS-FEC sublayer is shown in Figure 91–2.

91.5.2 Transmit function

91.5.2.1 Lane block synchronization

The RS-FEC transmit function forms 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 obtains lock to the 66-bit blocks in each bit stream using the sync headers and outputs 66-bit blocks. Block lock is obtained as specified in the block lock state diagram shown in Figure 82–10.

91.5.2.2 Alignment lock and deskew

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 82–11. This process identifies the PCS lane number 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 82–12. The RS-FEC transmit function shall support a maximum Skew of 29 ns between PCS lanes and a maximum Skew Variation of 0.2 ns. Skew and Skew Variation are defined in 80.5.

91.5.2.3 Lane reorder

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

91.5.2.4 Alignment marker removal

After all PCS lanes are aligned and deskewed, the PCS lanes are multiplexed together in the proper 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 am_lock is true for a given PCS lane even if it does not match the expected alignment marker value (due to a bit error for example). Repeated alignment marker errors will result in am_lock being set to false for a given PCS lane, but until that happens it is sufficient to remove the block in the alignment marker position.

As part of the alignment marker removal process, the BIP_3 field is compared to the calculated Bit Interleaved Parity (BIP) value (refer to 82.2.8) for each PCS lane. If a Clause 45 MDIO is implemented, then the appropriate BIP error counter register (registers TBD through TBD) is incremented by one each time the calculated BIP value does not equal the value received in the BIP_3 field. The incoming bit error ratio can be estimated by dividing the BIP block error ratio by a factor of 1081344.

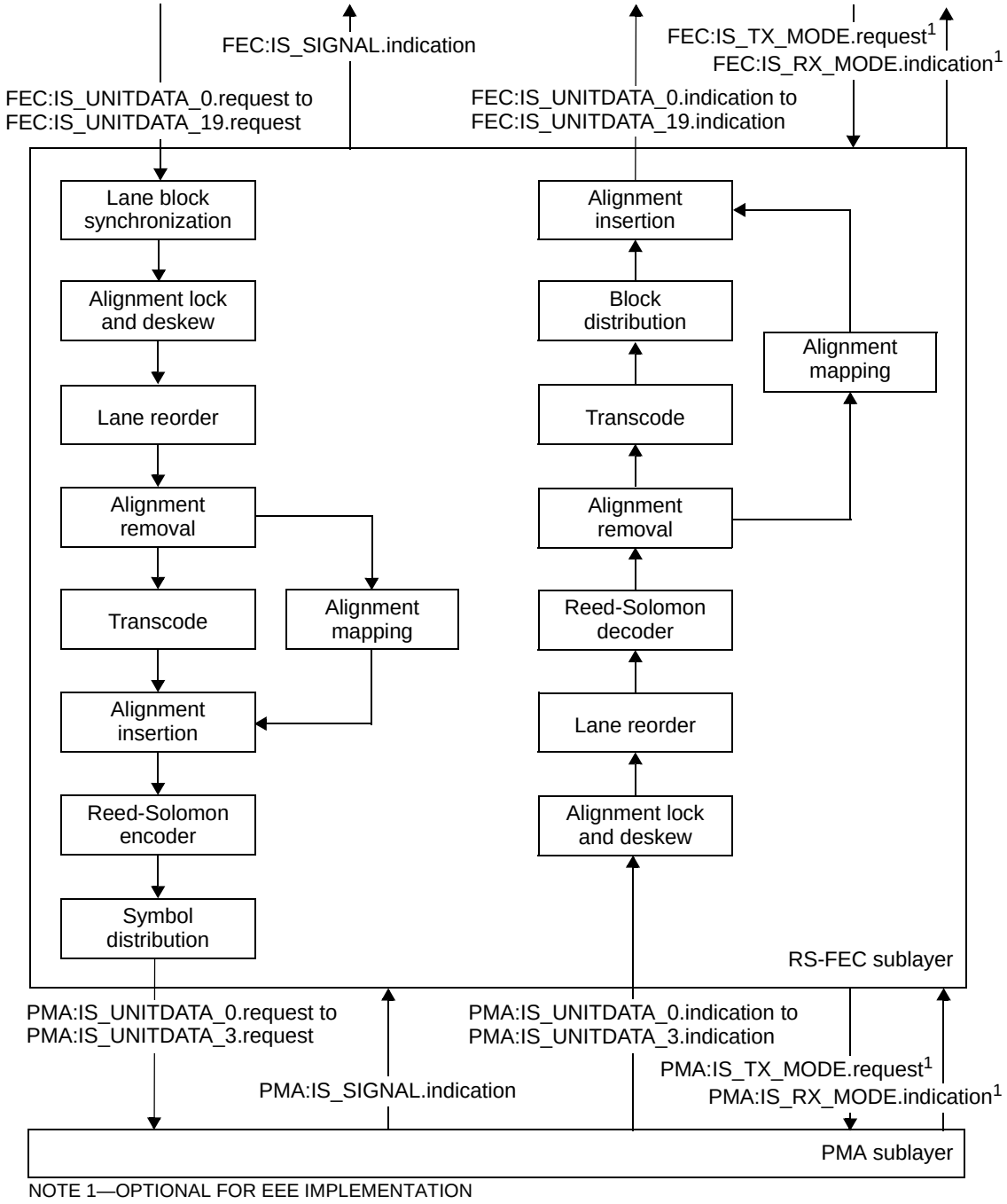


Figure 91–2—Functional block diagram

91.5.2.5 64B/66B to 256B/257B transcoder

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

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

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

If for any *j*=0 to 3, tx_coded_*j*<1>=0 and tx_coded_*j*<0>=1, 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*<1:0> = 01. In other words, tx_coded_*c* is the first 66-bit control block that was received in the current group of four blocks.
d) Let tx_payloads<(64*j*+63):64*j*> = 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.
   tx_xcoded<(64*c*+8):5> = tx_payloads<(64*c*+3):0>
   tx_xcoded<256:(64*c*+9)> = tx_payloads<255:(64*c*+8)>

Finally, tx_xcoded<4:0> shall be set to the result of the bit-wise exclusive-OR of the original tx_xcoded<4:0> and tx_xcoded<12:8>. Several examples that illustrate the transcoding process are shown in Figure 91–3.

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

91.5.2.6 Alignment mapping and insertion

The alignment markers that were removed per 91.5.2.4 are re-inserted after being processed by the alignment mapping function. The alignment mapping function compensates for the operation of the symbol distribution function defined in 91.5.2.8 and rearranges the alignment marker bits so that they appear on the FEC lanes intact and in the desired sequence. This preserves the properties of the alignment markers (e.g. DC balance, transition density) and provides a deterministic pattern for the purpose of synchronization.

The alignment mapping function operates on a group of twenty aligned and reordered alignment markers. Let am_*x*<65:0> be the alignment marker for PCS lane *x*, *x*=0 to 19, where bit 0 is the first bit transmitted. The alignment markers shall be mapped to am_txmapped<1284:0> in a manner that yields the same result as the process defined below.

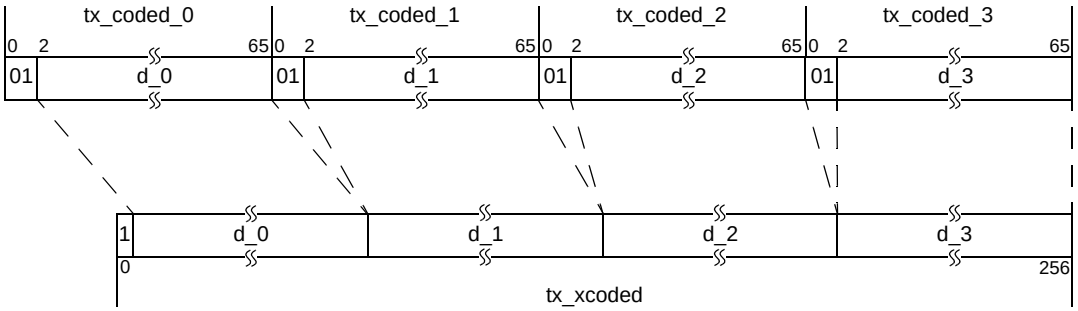
Discard the synchronization headers, am_*x*<1:0>, and construct a matrix of 4 rows and 320 columns, am_payloads, from the 64-bit payloads as shown in Figure 91–4. Given *i*=0 to 3, *j*=0 to 5, and *x*=*i*+4*j*, the matrix is defined by the following expression.

am_payloads<*i*, (64*j*+63):64*j*> = am_*x*<65:2>

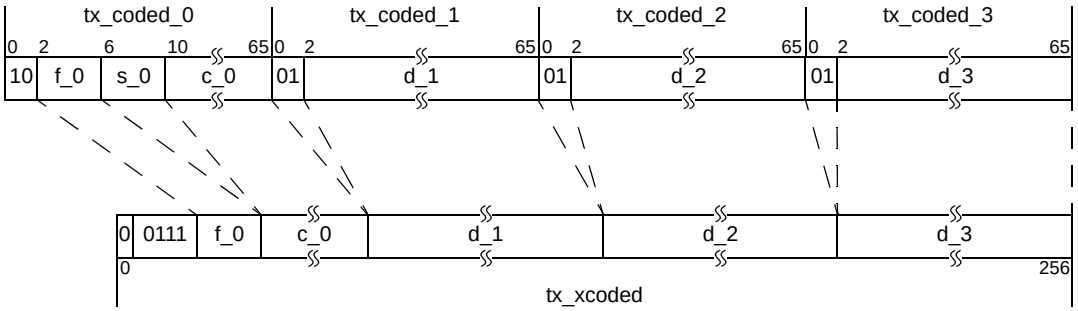
Given *i*=0 to 3, *k*=0 to 31, and *y*=*i*+4*k*, am_txmapped may then be derived from am_payloads per the following expression.

am_txmapped<(10*y*+9):10*y*> = am_payloads<*i*, (10*k*+9):10*k*>

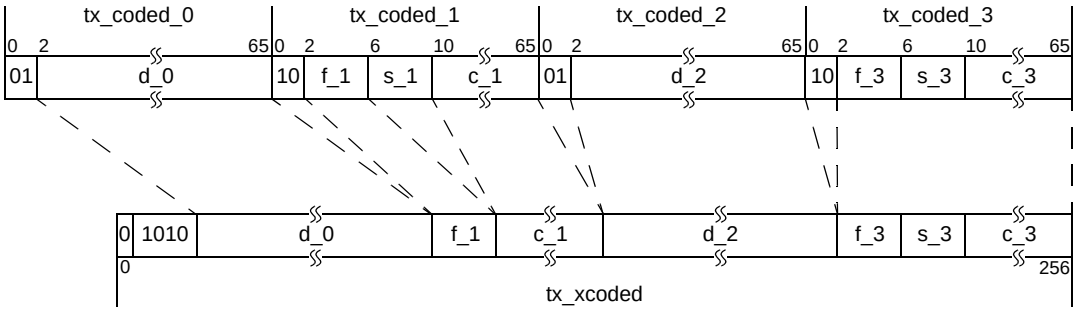
A 5-bit pad is appended to the mapped alignment markers to yield the equivalent of five 257-bit blocks. The pad bits, am_txmapped<1284:1280>, shall be set to 0x05 and 0x1A in an alternating pattern. In other words, if a pad value of 0x05 is used for the current iteration of the mapping function, a value of 0x1A will be used in the next iteration and vice versa.



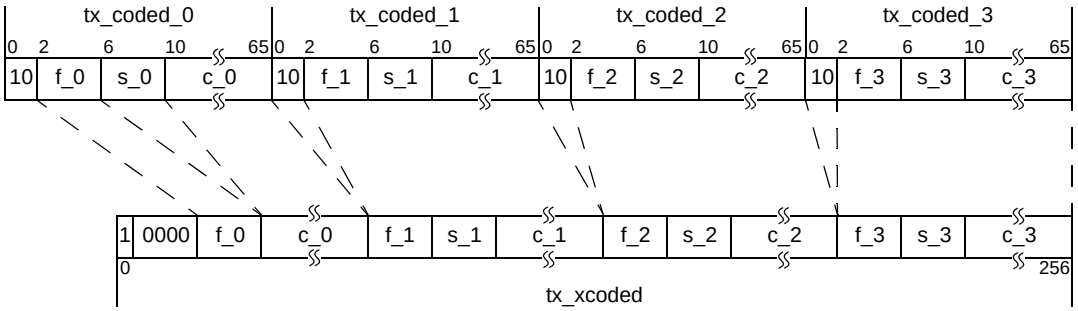
Example 1: All data blocks



Example 2: Control block followed by three data blocks



Example 3: Alternating data and control blocks



Example 4: All control blocks

Figure 91–3—64B/66B to 256B/257B transcoding examples

FEC lane, *i*	Reed-Solomon symbol index, *k* 0									31
0	$_2$	am_0 $_{65}$	$_2$	am_4 $_{65}$	$_2$	am_8 $_{65}$	$_2$	am_12 $_{65}$	$_2$	am_16 $_{65}$
1	$_2$	am_1 $_{65}$	$_2$	am_5 $_{65}$	$_2$	am_9 $_{65}$	$_2$	am_13 $_{65}$	$_2$	am_17 $_{65}$
2	$_2$	am_2 $_{65}$	$_2$	am_6 $_{65}$	$_2$	am_10 $_{65}$	$_2$	am_14 $_{65}$	$_2$	am_18 $_{65}$
3	$_2$	am_3 $_{65}$	$_2$	am_7 $_{65}$	$_2$	am_11 $_{65}$	$_2$	am_15 $_{65}$	$_2$	am_19 $_{65}$

Figure 91–4—Alignment marker mapping to FEC lanes

As a result of this process, the BIP_3 and BIP_7 fields from normal alignment markers are carried across the link protected by FEC. It should be noted that these fields cannot be used to monitor errors on the link protected by FEC as 64B/66B to 256B/257B trancoding and Reed-Solomon encoding alters the bit sequence. However, these fields may again be used to monitor errors after the original bit sequence is restored, i.e. following Reed-Solomon decoding and 256B/257B to 64B/66B transcoding.

One group of aligned and reordered alignment markers are mapped every 20 × 16384 66-bit blocks. This corresponds to 4096 Reed-Solomon codewords (refer to 91.5.2.7). The mapped alignment markers, am_txmapped<1284:0> shall be inserted as the first 1285 message bits to be transmitted from every 4096th codeword.

For the optional EEE capability, one group of Rapid Alignment Markers (see 82.2.7a) are mapped every 20 × 8 66-bit blocks. This corresponds to 2 Reed-Solomon codewords. The mapped Rapid Alignment Markers, am_txmapped<1284:0> shall be inserted as the first 1285 message bits to be transmitted from every other codeword.

91.5.2.7 Reed-Solomon encoder

The RS-FEC sublayer employs a Reed-Solomon code operating over the Galois Field GF(2^w) where w=10 is the symbol size in 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 this clause, a particular Reed-Solomon code is denoted RS(n, k).

When used to form a 100GBASE-CR4 or 100GBASE-KR4 PHY, the RS-FEC sublayer shall implement RS(528, 514). Each k-symbol message corresponds to twenty 257-bit blocks produced by the transcoder. This code has the capability to correct any combination of up to t=7 symbol errors in a codeword.

When used to form a 100GBASE-KP4 PHY, the RS-FEC sublayer shall implement RS(544, 514). Each k-symbol message corresponds to twenty 257-bit blocks produced by the transcoder. This code has the capability to correct any combination of up to t=15 symbol errors in a codeword.

Each code is based on the generating polynomial given by Equation (91–1).

$$g(x) = \prod_{j=0}^{2t-1} (x - \alpha^j) = g_{2t}x^{2t} + g_{2t-1}x^{2t-1} + \ldots + g_1 x + g_0 \quad (91–1)$$

In Equation (91–1), α is a primitive element of the finite field defined by the polynomial x^{10}+x^3+1.

Equation (91–2) defines the message polynominal $m(x)$ whose coefficients are the message symbols m_{k-1} to m_0.

$$m(x) = m_{k-1}x^{n-1} + m_{k-2}x^{n-2} + \ldots + m_1 x^{2t+1} + m_0 x^{2t} \quad (91\text{–}2)$$

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 composed of the bits of the transcoded blocks (including a 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 to the encoder is m_{k-1}.

Equation (91–3) defines the parity polynomial $p(x)$ whose the coefficients are the parity symbols p_{2t-1} to p_0.

$$p(x) = p_{2t-1}x^{2t-1} + p_{2t-2}x^{2t-2} + \ldots + p_1 x + p_0 \quad (91\text{–}3)$$

The parity polynomial is the remainder from the division of $m(x)$ by $g(x)$. This may be computed using the shift register implementation illustrated in Figure 91–5. The outputs of the delay elements are initialized to zero prior to the computation of the parity for a given message. After the last message symbol, m_0, is processed by the encoder, the outputs of the delay elements are the parity symbols for that message.

The codeword polynomial $c(x)$ is then 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 bit 0.

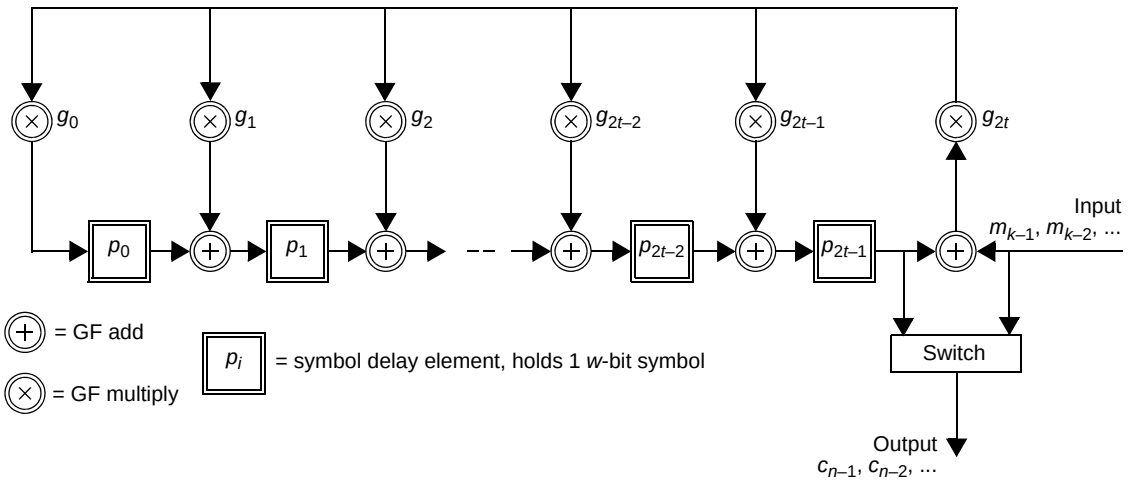


Figure 91–5—Reed-Solomon encoder functional model

The coefficients of the generator polynomial and example codewords for each code are described in Annex 91A.

Editor's note (to be removed prior to final publication):

Annex 91A will be added to include FEC codeword examples in the style of Annex 74A. The generation of these examples will be deferred until the content of this clause stabilizes.

91.5.2.8 Symbol distribution

Once the data is scrambled and encoded, it shall be distributed to 4 FEC lanes, one *w*-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 91–6.

When used to form a 100GBASE-KP4 PHY, the PMA:IS_UNITDATA_*i*.request primitive is defined to include an additional parameter. At the beginning of an FEC codeword, the parameter start=TRUE is asserted for the first bit of the first four symbols of the codeword transferred across the four primitives. Otherwise the parameter start is set to FALSE.

91.5.2.9 Transmit bit ordering

The transmit bit ordering is illustrated in Figure 91–6.

91.5.3 Receive function

91.5.3.1 Alignment lock and deskew

The RS-FEC receive function forms 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 91–8.

After alignment marker lock is achieved on all 4 lanes, all inter-lane Skew is removed as specified by the FEC deskew state diagram shown in Figure 91–9. The FEC receive function shall support a maximum Skew of 134 ns between FEC lanes and a maximum Skew Variation of 3.4 ns.

The 100GBASE-KP4 PMA transmit function (refer to 94.2.2) inserts PMA-specific overhead that is aligned with the start of a Reed-Solomon codeword. The 100GBASE-KP4 PMA receive function (refer to 94.2.4) synchronizes to this overhead and indicates the first bit of each of the first four symbols in a codeword by setting the PMA:IS_UNITDATA_*i*.indicate parameter start=TRUE.

91.5.3.2 Lane reorder

FEC lanes can be received on different lanes of the service interface from which they were originally transmitted due to connection errors in the underlying medium. The FEC receive function shall order the FEC lanes according to the FEC lane number. The FEC lane number is defined by the sequence of alignment markers that are mapped to each FEC lane.

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

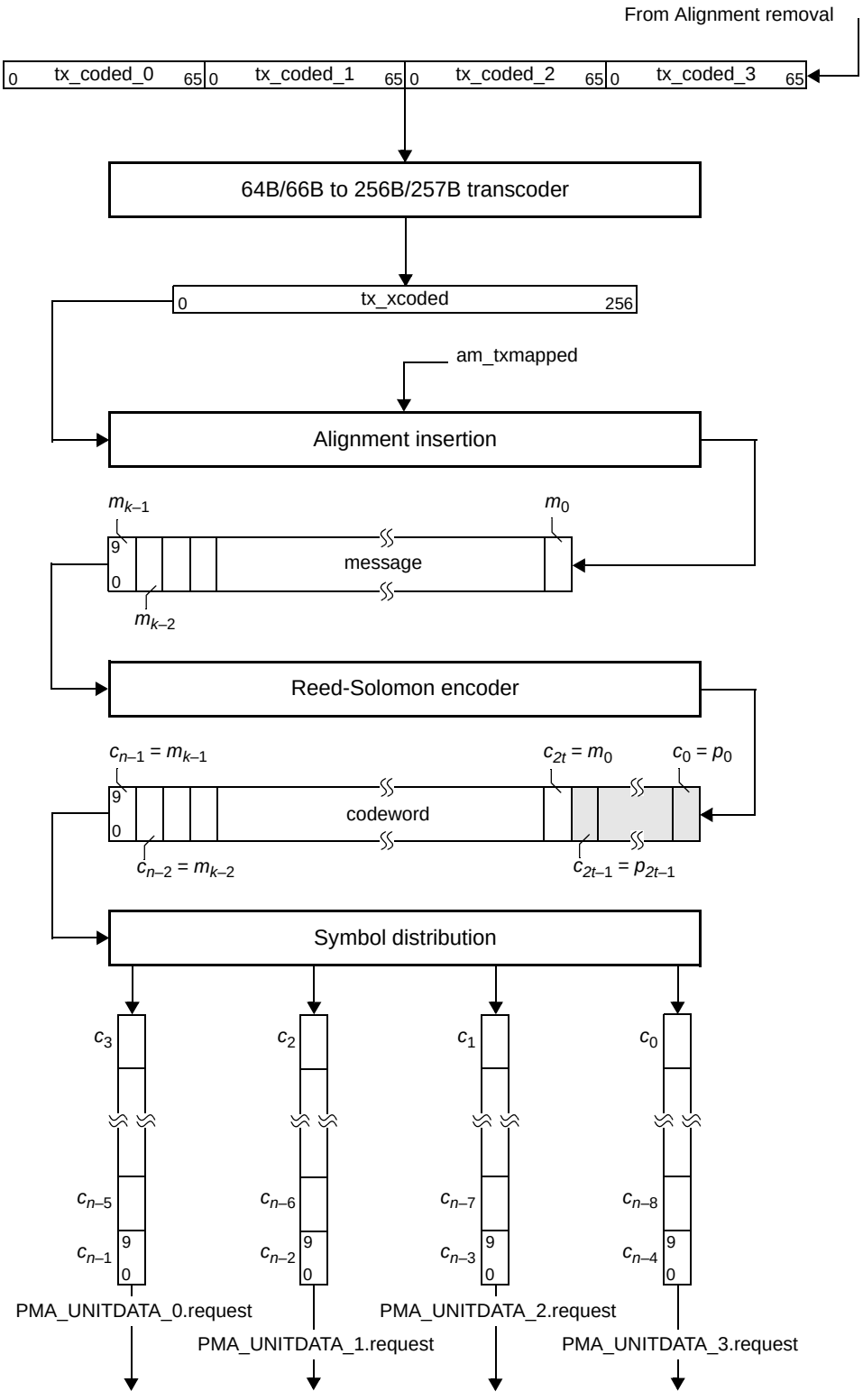


Figure 91–6—Transmit bit ordering

91.5.3.3 Reed-Solomon decoder

The Reed-Solomon decoder extracts the message symbols from the codeword, correcting them as necessary, and discards the parity symbols.

If the decoder determines that a codeword is uncorrectable, it shall ensure that, for every other 257-bit block within the codeword starting with the first (1st, 3rd, 5th, etc.), the synchronization header for the first 66-bit block at the output of the 256B/267B to 64B/66B transcoder, rx_coded_0<1:0>, is set to 11. In addition, it shall ensure rx_coded_3<1:0> corresponding to the last (20th) 257-bit block in the codeword is set to 11. This will cause all packets 64-bytes and larger within the uncorrectable codeword to be discarded by the PCS.

91.5.3.4 Alignment marker removal

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 include this vector are indicated by the alignment lock and deskew function (refer to 91.5.3.1). The vector am_rxmapped shall be removed prior to descrambling and transcoding.

91.5.3.5 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_xcoded<256:0>. Bit 0 of the 257-bit block is the first bit received.

First, rx_rxcoded<4:0> shall be set to the result of the bit-wise exclusive-OR of the received rx_xcoded<4:0> and rx_xcoded<12:8>.

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

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

a) if rx_xcoded<*j*+1> = 0, rx_coded_*j*<1> = 0 and rx_coded_*j*<0> = 1
b) if rx_xcoded<*j*+1> = 1, rx_coded_*j*<1> = 1 and rx_coded_*j*<0> = 0
c) Let *c* be the smallest value of *j* such that rx_xcoded<*j*+1> = 0. In other words, rx_coded_*c* will be the first 66-bit control block in the resulting group of four blocks.
d) Let rx_payloads be a vector representing the payloads of the four 66-bit blocks. It is derived using the following expressions.
   rx_payloads<(64*c*+3):0> = rx_xcoded<(64*c*+8):5>
   rx_payloads<(64*c*+7):(64*c*+4)> = 0000 (an arbitrary value that is later replaced, see step j)
   rx_payloads<255:(64*c*+8)> = rx_xcoded<256:(64*c*+9)>
e) rx_coded_*j*<65:2> = rx_payloads<(64*j*+63):64*j*> for *j*=0 to 3
f) 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*.
g) 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
h) 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 82–5. For example, if g<3:0> is 0xE then h<3:0> is 0x1.

i) Scramble h<3:0> to yield s_*c*<3:0> per the following expression.
s_*c*<*i*> = h<*i*> ^ rx_coded_(*c*–1)<*i*+12> ^ rx_coded_(*c*–1)<*i*+31> for *i*=0 to 3
j) 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.

91.5.3.6 Block distribution

Once the data is encoded and scrambled, it shall be distributed to multiple PCS lanes, 66-bit blocks at a time in a round robin distribution from the lowest to the highest numbered PCS lanes. The distribution process is shown in Figure 82–6.

91.5.3.7 Alignment marker mapping and insertion

The alignment marker mapping function compensates for operation of lane reorder function (refer to 91.5.3.2) to derive the PCS lane alignment markers, am_*x*<65:0> for *x*=0 to 19, from am_rxmapped<1284:0> (refer to 91.5.3.4).

The alignment markers shall be derived from am_rxmapped<1284:0> in a manner that yields the same result as the process defined below.

Given *i*=0 to 3, *k*=0 to 31, and *y*=*i*+4*k*, am_payloads may be derived from am_rxmapped per the following expression.

am_payloads<*i*, (10*k*+9):10*k*> = am_rxmapped<(10*y*+9):10*y*>

Given *i*=0 to 3, *j*=0 to 5, and *x*=*i*+4*j*, the am_*x* may be derived from am_payloads by the following expression.

am_*x*<65:2> = am_payloads<*i*, (64*j*+63):64*j*>

Finally, am_*x*<1:0> = 01 for all *x*=0 to 19.

91.5.3.8 Receive bit ordering

The receive bit ordering is illustrated in Figure 91–7. This illustration shows the case where the FEC lanes appear across the PMA:IS_UNITDATA_*i*.indication primitives in the correct order.

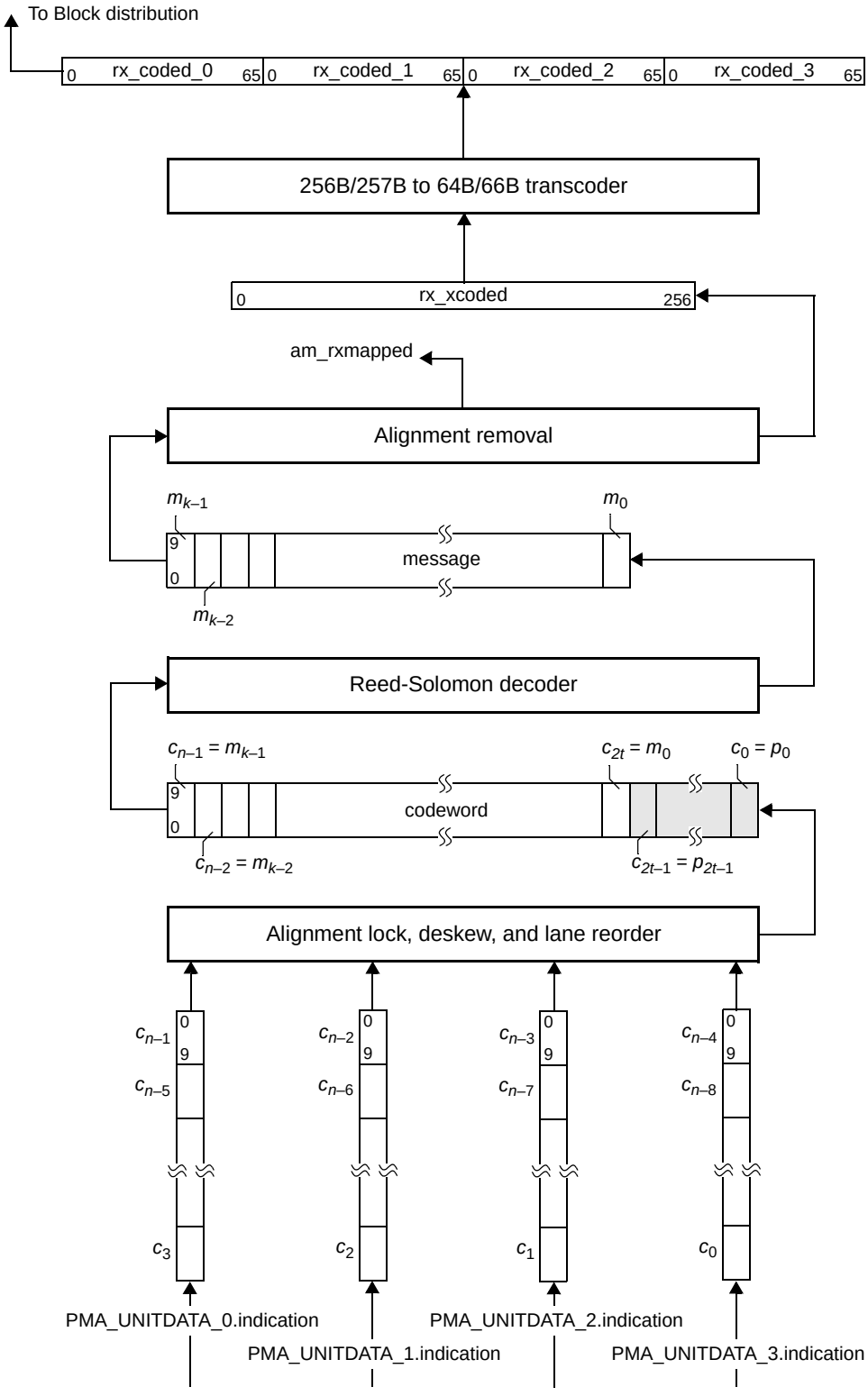


Figure 91–7—Receive bit ordering

91.5.4 Detailed functions and state diagrams

91.5.4.1 State diagram conventions

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

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

91.5.4.2 State variables

91.5.4.2.1 Variables

align_status
: A variable set by the FEC deskew process to reflect the status of FEC lane-to-lane alignment. Set to true when all lanes are synchronized and aligned and set to false when the deskew process is not complete.

alignment_valid
: Boolean variable that is set to true if all FEC lanes are aligned. FEC lanes are considered to be aligned when amps_lock<*x*> is true for all *x* and the FEC lanes are deskewed. Otherwise, this variable is set to false.

all_locked
: A Boolean variable that is set to true when amps_lock<*x*> is true for all *x* and is set to false when am_lock<*x*> is false for any *x*.

amp_counter_done
: Boolean variable that indicates that amp_counter has reached its terminal count.

amp_match
: Boolean variable that holds the output of the function AMP_COMPARE.

amp_valid
: Boolean variable that is set to true if the received 64-bit block is a valid alignment marker payload. The alignment marker payload, mapped to an FEC lane according to the process described in 91.5.2.6, consists of 48 known bits unique to each PCS lane and 16 variable bits (the BIP_3 or CD_3 field and its complement BIP_7 or CD_7, see 82.2.7). The bits of the candidate block that are in the positions of the known bits in the alignment marker payload are compared on a nibble-wise basis (12 comparisons). If no more than 3 nibbles in the candidate block fail to match the corresponding known nibbles in the alignment marker payload, the candidate block is considered a valid alignment marker payload. For the normal mode of operation, each FEC lane compares the candidate block to the alignment marker payload for PCS lanes 0, 1, 2, and 3. For the optional EEE capability, each FEC lane shall also compare the candidate block to the alignment marker payload for PCS lanes 16, 17, 18, and 19.

Editor's note (to be removed prior to final publication):

The maximum distance permitted between the candidate block and the reference pattern (3 nibbles) was calculated for the 100GBASE-KR4 PHY. It should be confirmed that this value is also suitable for the 100GBASE-KP4 PHY.

amps_lock<*x*>

Boolean variable that is set to true when the receiver has detected the location of the alignment maker payload sequence for a given FEC lane where *x* = 0:3.

current_pcsl

A variable that holds the PCS lane number corresponding to the current alignment marker payload that is recognized on a given FEC lane. It is compared to the variable first_pscl to confirm that the location of the alignment marker payload sequence has been detected.

cw_bad

A Boolean variable that is set to true if the Reed-Solomon decoder (see 91.5.3.3) determines that the current FEC codeword is uncorrectable and is set to false otherwise.

enable_deskew

A Boolean variable that enables and disables the deskew process. Received bits may be discarded whenever deskew is enabled. It is set to true when deskew is enabled and it is set to false when deskew is disabled.

first_pcsl

A variable that holds the PCS lane number that corresponds to the first alignment marker payload that is recognized on a given FEC lane. It is compared to the PCS lane number corresponding to the second alignment maker payload that is tested.

reset

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

signal_ok

Boolean variable that is set based on the most recently received value of *inst*:IS_SIGNAL.indication(SIGNAL_OK). It is true if the value was OK and false if the value was FAIL.

slip_done

Boolean variable that is set to true when the SLIP requested by the synchronization state diagram has been completed indicating that the next candidate 64-bit block position can be tested.

test_amp

Boolean variable this is set to true when a candidate block position is available for testing and false when the FIND_1ST state is entered.

test_cw

Boolean variable that is set to true when a new FEC codeword is available for decoding and is set to false when the TEST_CW state is entered.

91.5.4.2.2 Functions

AMP_COMPARE

This function compares the values of first_pcsl and current_pcsl to determine if a valid alignment marker payload sequence has been detected and returns the result of the comparison using the variable amp_match. In the normal mode of operation, if current_pcsl equals first_pcsl, amp_match is set to true. For the optional EEE capability, if first_pcsl is either 0, 1, 2, or 3 then amp_match is set to true if current_pcsl is first_pcsl+16. Also for the optional EEE capability, if first_pcsl is 16, 17, 18, or 19 then amp_match is set to true if current_pcsl is first_pcsl–16. Otherwise, amp_match is set to false.

SLIP

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

91.5.4.2.3 Counters

amp_counter

In the normal mode, this counter counts the 4096 FEC codewords that separate the ends of two

consecutive alignment marker payload sequences. Note that an FEC codeword is 1320 bits per FEC lane for 100GBASE-KR4 and 1360 bits per FEC lane for 100GBASE-KP4. For the optional EEE capability, if first_amp corresponds to PCS lane 0, 1, 2, or 3, it counts the 256 bits to the end of the expected location of the alignment marker payload corresponding to PCS lanes 16, 17, 18, or 19. Also for the optional EEE capability, if first_amp corresponds to PCS lane 16, 17, 18, or 19, this counter counts the 4096 FEC codewords minus 256 bits to the end of the expected location of the next alignment marker payload corresponding to PCS lanes 0, 1, 2, or 3.

cw_bad_count

Counts the number of consecutive uncorrectable FEC codewords. This counter is set to zero when an FEC codeword is received and cw_bad is false for that codeword.

91.5.4.3 State diagrams

The FEC shall implement four synchronization processes as shown in Figure 91–8. The synchronization process operates independently on each lane. The synchronization state diagram shown determines when the FEC has detected the location of the alignment marker payload sequence in the received bit stream for a given lane of the service interface.

The FEC shall implement the deskew process as shown in Figure 91–9.

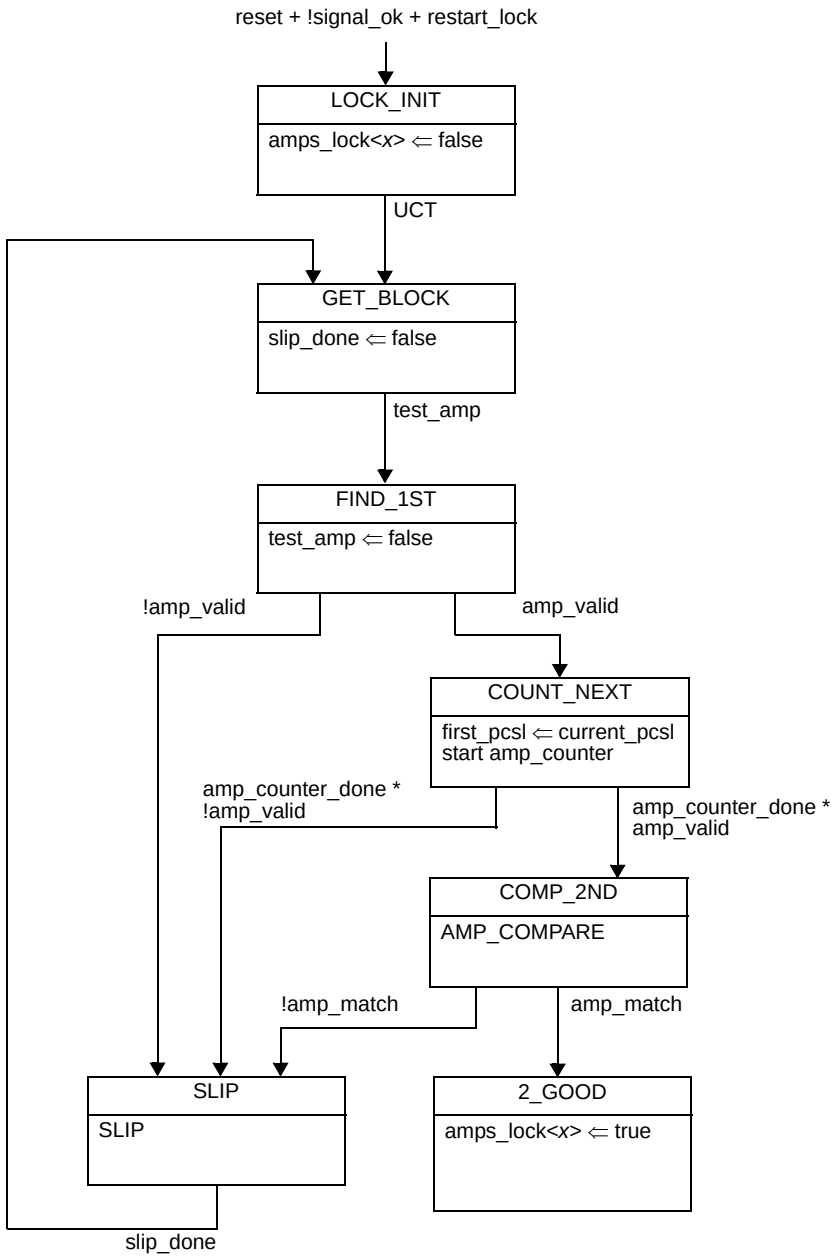


Figure 91–8—FEC synchronization state diagram

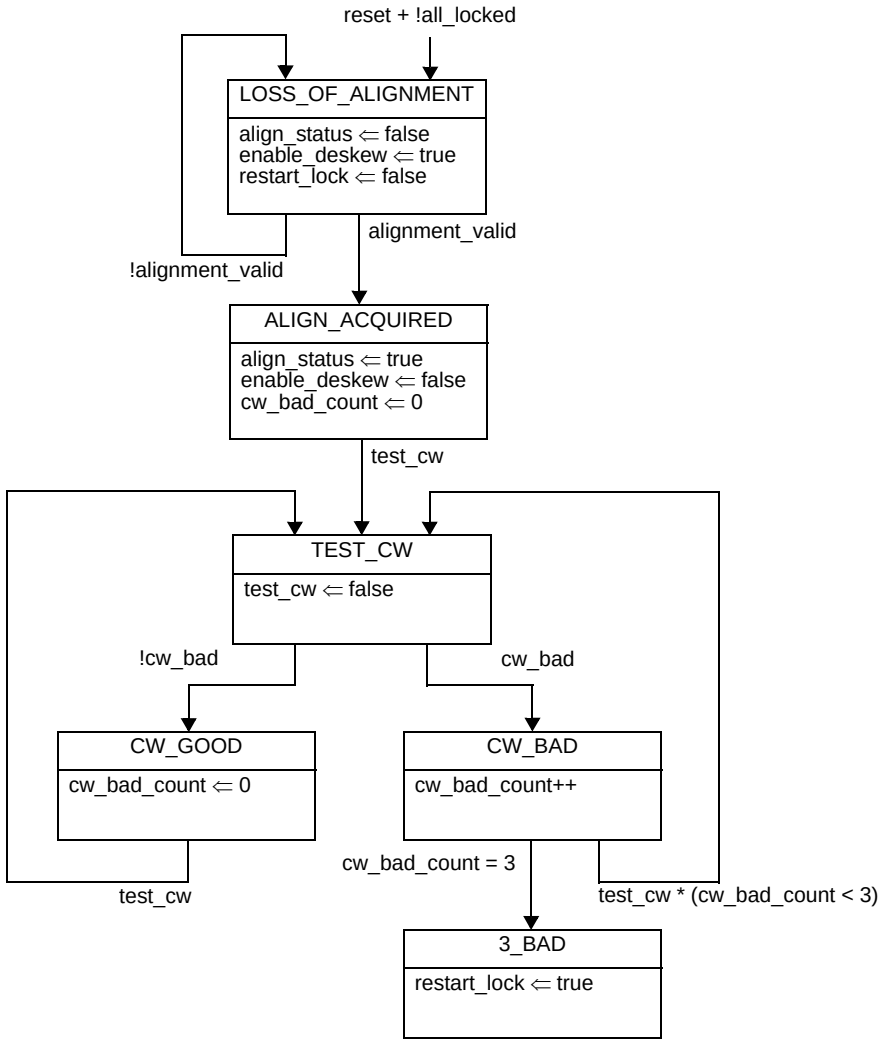


Figure 91–9—FEC deskew state diagram

91.6 FEC management

Editor's note (to be removed prior to final publication):

The adopted baselines did not address RS-FEC management. The following text has been inserted as a placeholder.

The optional MDIO capability described in Clause 45 defines several variables that provide control and status information for and about the FEC. If MDIO is implemented, it shall map MDIO control variables to FEC control variables as shown in Table 91–X, and MDIO status variables to FEC status variables as shown in Table 91–X.

91.7 Protocol implementation conformance statement (PICS) proforma for Clause 91, Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs[9]

91.7.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 91, Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs, shall complete the following protocol implementation conformance statement (PICS) proforma.

A detailed description of the symbols used in the PICS proforma, along with instructions for completing the PICS proforma, can be found in Clause 21.

91.7.2 Identification

91.7.2.1 Implementation identification

Supplier[1]	
Contact point for enquiries about the PICS[1]	
Implementation Name(s) and Version(s)[1,3]	
Other information necessary for full identification—e.g., name(s) and version(s) for machines and/or operating systems; System Name(s)[2]	
NOTE 1— Required for all implementations. NOTE 2— May be completed as appropriate in meeting the requirements for the identification. NOTE 3—The terms Name and Version should be interpreted appropriately to correspond with a supplier's terminology (e.g., Type, Series, Model).	

91.7.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3bj-20XX, Clause 91, Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs
Identification of amendments and corrigenda to this PICS proforma that have been completed as part of this PICS	
Have any Exception items been required? No [ ] Yes [ ] (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3bj-20XX.)	

Date of Statement	

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

Major capabilities/options

Editor's note (to be removed prior to final publication):

Major capabilities/options and PICS proforma tables will be generated when the content of the clause stabilizes.

92. Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-CR4

92.1 Overview

This clause specifies the 100GBASE-CR4 PMD and baseband medium. Annex 92A, an associated annex, provides information on parameters with test points that may not be testable in an implemented system. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 92–1, to the medium through the MDI and to the management functions that are optionally accessible through the management interface defined in Clause 45, or equivalent.

When forming a complete 100GBASE-CR4 Physical Layer, the following guidelines apply.

a) The PMA specified in Clause 83 shall connect to the RS-FEC specified in Clause 91 using the PMA service interface specified in 83.3 with p=4 input lanes.
b) The PMA specified in Clause 83 shall connect to the PMD using the service interface specified in 92.2.
c) If one or more CAUI interfaces are implemented between the RS and the RS-FEC, each CAUI interface shall include a pair of PMA sub-layers, as specified in Clause 83. Refer to 83.1.4 for additional guidance.

Table 92–1—Physical Layer clauses associated with the 100GBASE-CR4 PMD

Associated clause	100GBASE-CR4
81—RS	Required
81—CGMII[a]	Optional
82—PCS for 100GBASE-R	Required
91—RS-FEC	Required
83—PMA for 100GBASE-R	Required
83A—CAUI	Optional
73—Auto-Negotiation	Required
78—Energy Efficient Ethernet	Optional

[a]The CGMII is an optional interface. However, if the CGMII is not implemented, a conforming implementation must behave functionally as though the RS and CGMII were present.

A 100GBASE-CR4 PHY with the optional Energy-Efficient Ethernet (EEE) capability may optionally enter the Low Power Idle (LPI) mode to conserve energy during periods of low link utilization.

Figure 92–1 shows the relationship of the 100GBASE-CR4 PMD sublayers and MDI to the ISO/IEC Open System Interconnection (OSI) reference model.

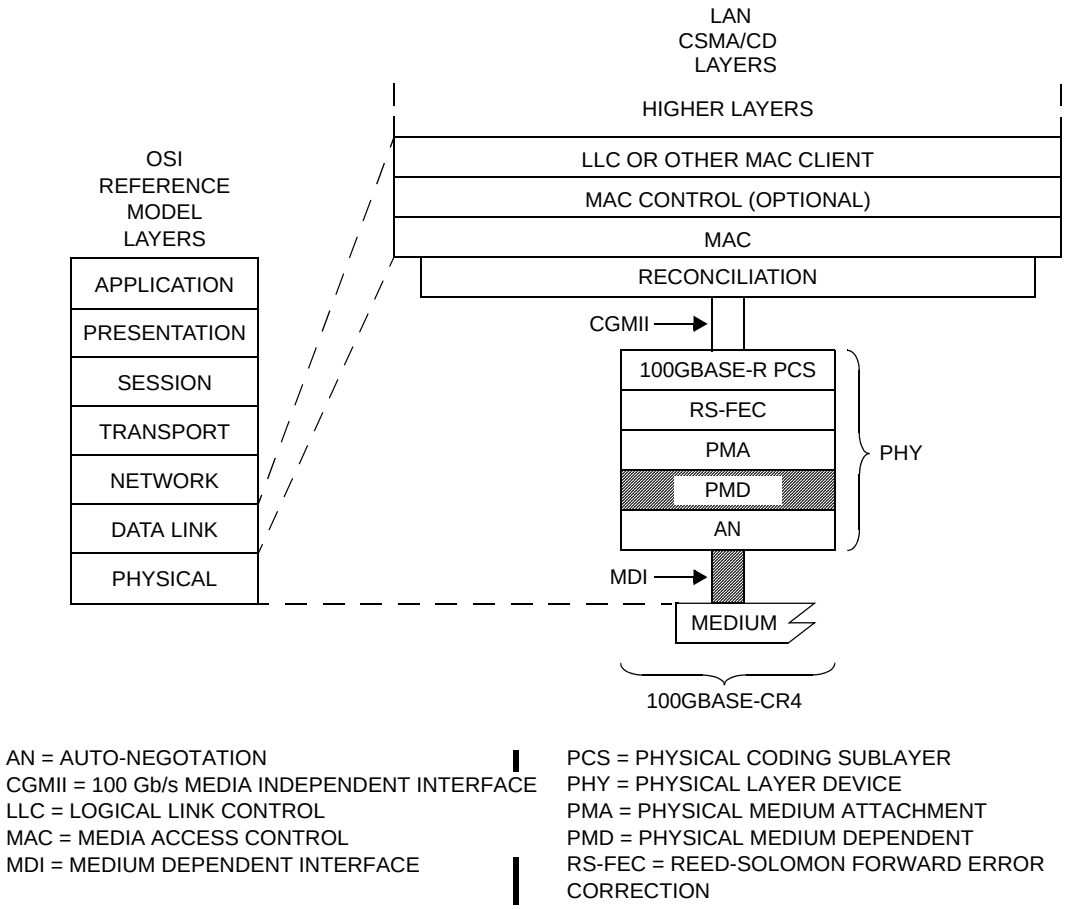


Figure 92–1—100GBASE-CR4 PMD relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

92.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 100GBASE-CR4 PMD. The service interface for this PMD is described in an abstract manner and does not imply any particular implementation. The PMD service interface supports the exchange of encoded data. The PMD translates the encoded data to and from signals suitable for the medium.

The PMD service interface is an instance of the inter-sublayer service interface defined in 80.3. The PMD service interface primitives are summarized as follows:

PMD:IS_UNITDATA_*i*.request
PMD:IS_UNITDATA_*i*.indication
PMD:IS_SIGNAL.indication

The 100GBASE-CR4 PMD has four parallel bit streams, hence $i = 0$ to 3. The PMA (or the PMD) continuously sends four parallel bit streams to the PMD (or the PMA), one per lane, each at a nominal signaling rate of 25.78125 GBd.

SIGNAL_DETECT in 100GBASE-CR4 indicates the successful completion of the start-up protocol on all four lanes. The SIGNAL_DETECT parameter can take on one of two values: OK or FAIL. When SIGNAL_DETECT = FAIL, the PMD:IS_UNITDATA_*i*.indication parameters are undefined. The SIGNAL_DETECT parameter maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

If the optional EEE capability is supported, then the PMD service interface includes two additional primitives as follows:

PMD:IS_TX_MODE.request
PMD:IS_RX_MODE.indicate

Editor's note (to be removed prior to final publication):

The PMD service interface has been augmented as implied by barrass_01_0312 per the motion passed in the March Task Force meeting.

"Move that the 802.3bj Task Force adopt barrass_01_0312.pdf as a baseline for optional Energy Efficient Ethernet operation for 100G Backplane and Twinaxial cable PHYs. (M: M. Gustlin, S: M. Brown, Y: 54, N: 0, A: 3)"

Per barrass_01_0312, PMD transmit function behavior changes as a function of the tx_mode parameter.

DATA, SLEEP, or WAKE: normal behavior
ALERT: send the ALERT signal
FW: send a PMA-specific pattern that is TBD (PMD normal behavior?)
QUIET: disable the transmitter

Per barrass_01_0312, the PMD receive function infers rx_mode from the received signal.

Normal or rapid alignment markers: DATA, SLEEP, or WAKE (mapping not defined)
No signal: QUIET
Alert signal: ALERT
PMA-specific pattern for FW that is TBD: FW (PMA function?)

Given the ambiguity surrounding PMD behavior and the TBD signaling for ALERT, FW, etc., these functions have not been added to the clause. The functions will be added to the clause upon adoption of a detailed proposal describing the desired behavior.

92.3 PCS requirements for Auto-Negotiation (AN) service interface

The PCS associated with this PMD is required to support the AN service interface primitive AN_LINK.indication defined in 73.9. (See 82.6.)

92.4 Delay constraints

The sum of the transmit and the receive delays at one end of the link contributed by the 100GBASE-CR4 PMD, AN, and the medium in one direction shall be no more than TBD bit times (TBD pause_quanta or TBD ns). It is assumed that the one way delay through the medium is no more than TBD bit times (TBD ns).

A description of overall system delay constraints and the definitions for bit times and pause_quanta can be found in 80.4 and its references.

92.5 Skew constraints

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. 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 80.5 and specified at the points SP1 to SP6 shown in Figure 80–4 and Figure 80–5.

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 TBD ns and the Skew Variation at SP2 is limited to TBD ps.

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

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

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 TBD ns and the Skew Variation at SP5 shall be less than TBD ns.

For more information on Skew and Skew Variation see 80.5.

92.6 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that provide control and status information for and about the PMD. If MDIO is implemented, it shall map MDIO control variables to PMD control variables as shown in Table 92–2, and MDIO status variables to PMD status variables as shown in Table 92–3.

Table 92–2—100GBASE-CR4 MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable	1.9.0	Global_PMD_transmit_disable
PMD transmit disable 3 to PMD transmit disable 0	PMD transmit disable	1.9.4 to 1.9.1	PMD_transmit_disable_3 to PMD_transmit_disable_0
Restart training	BASE-R PMD control	1.150.0	mr_restart_training
Training enable	BASE-R PMD control	1.150.1	mr_training_enable

92.7 PMD functional specifications

92.7.1 Link block diagram

A 100GBASE-CR4 link in one direction is illustrated in Figure 92–2. For purposes of system conformance, the PMD sublayer is standardized at the test points described in this subclause. The electrical transmit signal is defined at TP2. Unless specified otherwise, all transmitter measurements and tests defined in Table 92–5 are made at TP2 utilizing the test fixture specified in 92.8.3.5. Unless specified otherwise, all receiver

Table 92–3—100GBASE-CR4 MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMD status variable
Fault	PMA/PMD status 1	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect	1.10.0	Global_PMD_signal_detect
PMD receive signal detect 3 to PMD receive signal detect 0	PMD receive signal detect	1.10.4 to 1.10.1	PMD_signal_detect_9 to PMD_signal_detect_0
Receiver status 3	BASE-R PMD status	1.151.12	rx_trained_3
Frame lock 3	BASE-R PMD status	1.151.13	frame_lock_3
Start-up protocol status 3	BASE-R PMD status	1.151.14	training_3
Training failure 3	BASE-R PMD status	1.151.15	training_failure_3
Receiver status 2	BASE-R PMD status	1.151.8	rx_trained_2
Frame lock 2	BASE-R PMD status	1.151.9	frame_lock_2
Start-up protocol status 2	BASE-R PMD status	1.151.10	training_2
Training failure 2	BASE-R PMD status	1.151.11	training_failure_2
Receiver status 1	BASE-R PMD status	1.151.4	rx_trained_1
Frame lock 1	BASE-R PMD status	1.151.5	frame_lock_1
Start-up protocol status 1	BASE-R PMD status	1.151.6	training_1
Training failure 1	BASE-R PMD status	1.151.7	training_failure_1
Receiver status 0	BASE-R PMD status	1.151.0	rx_trained_0
Frame lock 0	BASE-R PMD status	1.151.1	frame_lock_0
Start-up protocol status 0	BASE-R PMD status	1.151.2	training_0
Training failure 0	BASE-R PMD status	1.151.3	training_failure_0

measurements and tests defined in 92.8.4 are made at TP3 utilizing the test fixture specified in 92.8.3.5. A mated connector pair has been included in both the transmitter and receiver specifications defined in 92.8.3 and 92.8.4. The maximum insertion loss from TP0 to TP2 or TP3 to TP5 including the test fixture is specified in 92.8.3.4.

The 100GBASE-CR4 channel is defined between the transmitter (TP0) and receiver blocks (TP5) to include the transmitter and receiver differential controlled impedance printed circuit board insertion loss and the cable assembly insertion loss as illustrated in Figure 92–2. Annex 92A provides information on parameters associated with test points TP0 and TP5 that may not be testable in an implemented system. All cable assembly measurements are to be made between TP1 and TP4 as illustrated in Figure 92–2. The cable assembly test fixture of Figure 92–12 or its equivalent, is required for measuring the cable assembly specifications in 92.10 at TP1 and TP4. Two mated connector pairs and the cable assembly test fixture have been included in the cable assembly specifications defined in 92.10. Transmitter and receiver differential controlled impedance printed circuit board insertion losses defined between TP0–TP1 and TP4–TP5

respectively are provided informatively in Annex 92A.

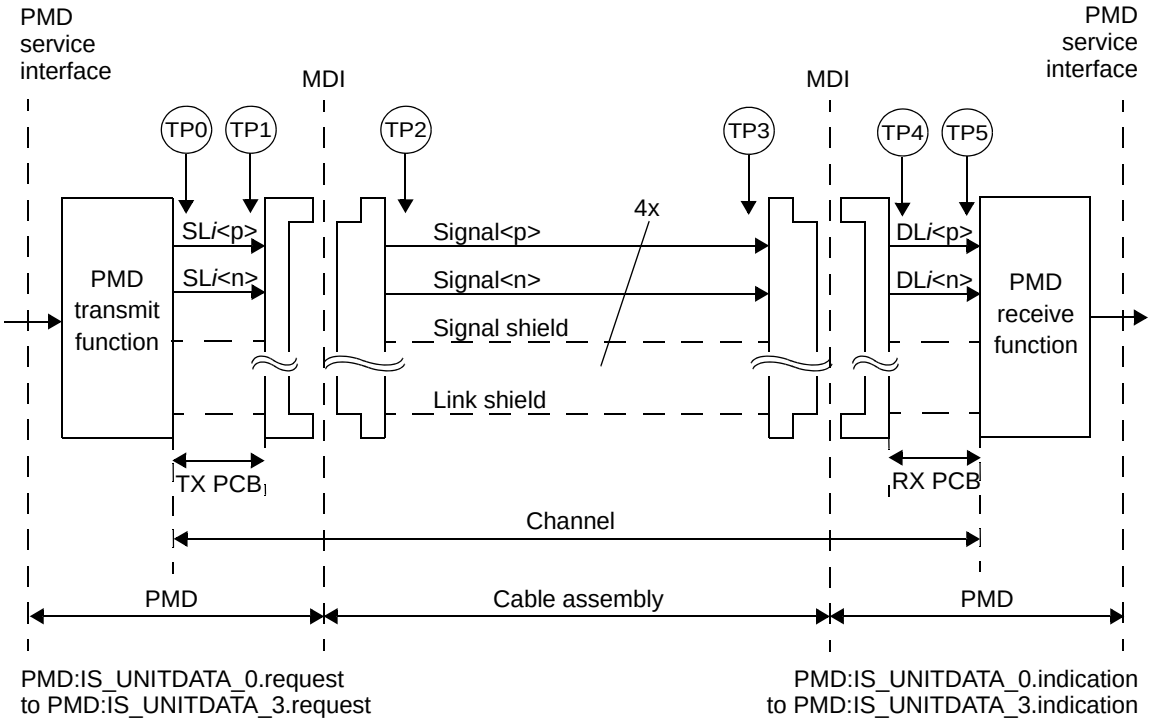


Figure 92–2—100GBASE-CR4 link (one direction is illustrated)

Note that the source lanes (SL), signals SLn<p>, and SLn<n> are the positive and negative sides of the transmitters differential signal pairs and the destination lanes (DL) signals, DLn<p>, and DLn<n> are the positive and negative sides of the receivers differential signal pairs for lane n (n = 0, 1, 2, 3)

Table 92–4 describes the defined test points illustrated in Figure 92–2.

Table 92–4—100GBASE-CR4 test points

Test points	Description
TP0 to TP5	The 100GBASE-CR4 channel including the transmitter and receiver differential controlled impedance printed circuit board insertion loss and the cable assembly insertion loss.
TP1 to TP4	All cable assembly measurements are to be made between TP1 and TP4 as illustrated in Figure 92–2. The cable assembly test fixture of Figure 92–12 or its equivalent, is required for measuring the cable assembly specifications in 92.10 at TP1 and TP4.
TP0 to TP1 TP4 to TP5	A mated connector pair has been included in both the transmitter and receiver specifications defined in 92.8.3 and 92.8.4. The maximum insertion loss from TP0 to TP2 or TP3 to TP5 including the test fixture is specified in 92.8.3.4.
TP2	Unless specified otherwise, all transmitter measurements and tests defined in Table 92–5 are made at TP2 utilizing the test fixture specified in 92.8.3.5.
TP3	Unless specified otherwise, all receiver measurements and tests defined in 92.8.4 are made at TP3 utilizing the test fixture specified in 92.8.3.5.

92.7.2 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 electrical streams. The four electrical signal streams shall then be delivered to the MDI, all according to the transmit electrical specifications in 92.8.3. A positive output voltage of SL*i*<p> minus SL*i*<n> (differential voltage) shall correspond to tx_bit = one.

92.7.3 PMD Receive function

The PMD receive function shall convert the four electrical streams from the MDI 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. A positive input voltage of DL*i*<p> minus DL*i*<n> (differential voltage) shall correspond to rx_bit = one.

92.7.4 Global PMD signal detect function

The Global PMD signal detect function shall continuously report the message PMD:IS_SIGNAL.indication (SIGNAL_DETECT) to the PMD service interface. SIGNAL_DETECT, while normally intended to be an indicator of signal presence, is used by the 100GBASE-CR4 PMD to indicate the successful completion of the start-up protocol on all lanes. The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

SIGNAL_DETECT shall be set to FAIL following system reset or the manual reset of the training state diagram. Upon successful completion of training on all lanes, SIGNAL_DETECT shall be set to OK.

If training is disabled by management, SIGNAL_DETECT shall be set to OK.

If the MDIO interface is implemented, then Global_PMD_signal_detect (1.10.0) shall be continuously set to the value of SIGNAL_DETECT as described in 45.2.1.9.7.

92.7.5 PMD lane-by-lane signal detect function

When the MDIO is implemented, each PMD_signal_detect_*i* value, where *i* represents the lane number in the range 0 to 3, shall be continuously updated as described in the following two paragraphs.

PMD_signal_detect_0 (1.10.1), PMD_signal_detect_1 (1.10.2), PMD_signal_detect_2 (1.10.3), and PMD_signal_detect_3 (1.10.4) shall be set to one or zero depending on whether a particular lane's signal_detect, as defined by the training state diagram in Figure 72–5, returns true or false.

92.7.6 Global PMD transmit disable function

The Global_PMD_transmit_disable function is optional. When implemented, it allows all of the transmitters to be disabled with a single variable.

d) When Global_PMD_transmit_disable variable is set to one, this function shall turn off all of the transmitters such that each transmitter drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage in Table 92–5.
e) If a PMD_fault (92.7.9) is detected, then the PMD may turn off the electrical transmitter in all lanes.
f) Loopback, as defined in 92.7.8, shall not be affected by Global_PMD_transmit_disable.

92.7.7 PMD lane-by-lane transmit disable function

The PMD_transmit_disable_*i* function (where *i* represents the lane number in the range 0:3) is optional and allows the electrical transmitter in each lane to be selectively disabled. When this function is supported, it shall meet the following requirements:

a) When a PMD_transmit_disable_*i* variable is set to one, this function shall turn off the transmitter associated with that variable such that it drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage specified in Table 92–5.
b) If a PMD_fault (92.7.9) is detected, then the PMD may set each PMD_transmit_disable_*i* to one, turning off the electrical transmitter in each lane.
c) Loopback, as defined in 92.7.8, shall not be affected by PMD_transmit_disable_*i*.

92.7.8 Loopback mode

Local loopback mode shall be provided by the adjacent PMA (see 83.5.8) as a test function to the device. When loopback mode is enabled, transmission requests passed to each transmitter are sent directly to the corresponding receiver, overriding any signal detected by each receiver on its attached link. Note that loopback mode does not affect the state of the transmitter, which continues to send data (unless disabled).

Control of the loopback function is specified in 45.2.1.1.5.

NOTE 1—The signal path that is exercised in the loopback mode is implementation specific, but it is recommended that this signal path encompass as much of the circuitry as is practical. The intention of providing this loopback mode of operation is to permit diagnostic or self-test functions to test the transmit and receive data paths using actual data. Other loopback signal paths may also be enabled independently using loopback controls within other devices or sublayers.

NOTE 2—Placing a network port into loopback mode can be disruptive to a network.

92.7.9 PMD_fault function

If the MDIO is implemented, PMD_fault is the logical OR of PMD_receive_fault, PMD_transmit_fault, and any other implementation specific fault.

92.7.10 PMD transmit fault function

The PMD_transmit_fault function is optional. The faults detected by this function are implementation specific, but should not include the assertion of the Global_PMD_transmit_disable function.

If a PMD_transmit_fault (optional) is detected, then the Global_PMD_transmit_disable function should also be asserted.

If the MDIO interface is implemented, then this function shall be mapped to the Transmit fault bit as specified in 45.2.1.7.4.

92.7.11 PMD receive fault function

The PMD_receive_fault function is optional. The faults detected by this function are implementation specific.

If the MDIO interface is implemented, then this function shall contribute to the Receive fault bit as specified in 45.2.1.7.5.

92.7.12 PMD control function

Each lane of the 100GBASE-CR4 PMD shall use the same control function as 10GBASE-KR, as defined in 72.6.10.

The random seed for the training pattern described in 72.6.10.2.6 shall be different for each of the lanes.

The variables rx_trained_*i*, frame_lock_*i*, training_*i*, and training_failure_*i* (where *i* goes from 0 to 3) report status for each lane and are equivalent to rx_trained, frame_lock, training, and training_failure as defined in 72.6.10.3.1.

If the MDIO interface is implemented, then this function shall map these variables to the appropriate bits in the BASE-R PMD status register (Register 1.151) as specified in 45.2.1.80.

92.8 100GBASE-CR4 electrical characteristics

92.8.1 Signal levels

The 100GBASE-CR4 MDI is a low-swing AC coupled differential interface. AC coupling at the receiver, as defined in 92.8.4.5, allows for interoperability between components operating from different supply voltages. Low-swing differential signaling provides noise immunity and improved electromagnetic interference (EMI).

92.8.2 Signal paths

The 100GBASE-CR4 MDI signal paths are point-to-point connections. Each path corresponds to a 100GBASE-CR4 MDI lane and comprises two complementary signals, which form a balanced differential pair. For 100GBASE-CR4, there are four differential paths in each direction for a total of eight pairs, or sixteen connections. The signal paths are intended to operate on twinaxial cable assemblies ranging from 0.5 m to 5 m in length, as described in 92.10.

92.8.3 Transmitter characteristics

Transmitter characteristics shall meet specifications summarized in Table 92–5 at TP2 unless otherwise noted. The transmitter specifications at TP0 are provided informatively in Annex 92A.

Table 92–5—Transmitter characteristics at TP2 summary

Parameter	Subclause reference	Value	Units
Signaling rate, per lane	92.8.3.9	25.78125±100 ppm	GBd
Unit interval nominal	92.8.3.9	38.787879	ps
Differential peak-to-peak output voltage (max) with Tx disabled		35	mV
Common-mode voltage limits	72.7.1.4	1.9	V
Differential output return loss (min)	92.8.3.1	See Equation (92–1)	dB
Common-mode AC output voltage (max., RMS)		30	mV
Amplitude peak-to-peak (max)	72.7.1.4	1200[a]	mV
Transmitter Steady state voltage[b]	92.8.3.3	0.34 min, 0.6 max	V
Linear fit pulse (min)[c]	92.8.3.3	0.52 x Transmitter Steady state voltage	V
Transmitted waveform max RMS normalized error (linear fit), "*e*" abs coefficient step size minimum precursor fullscale range minimum post cursor fullscale range	 92.8.3.3 92.8.3.3.2 92.8.3.3.3 92.8.3.3.3	 0.037 0.0083 min, 0.05 max 1.54 4	
Far-end transmit output noise (max) Low insertion loss channel High insertion loss channel	92.8.3.2	 2 Equation (92–2) 1Equation (92–3)	mV
Max output jitter (peak-to-peak) Effective random jitter Even-odd jitter Total jitter excluding data dependent jitter	92.8.3.8	 0.15 0.035 0.28	 UI UI UI

[a]The 100GBASE-CR4 Style-1 connector may support 100GBASE-CR4 or XLPPI interfaces. For implementations that support both interfaces, the transmitter should not exceed the XLPPI voltage maximum until a 100GBASE-CR4 cable assembly has been identified.
[b]The transmitter Steady state voltage is the sum of linear fit pulse response *p*(k) from step 3) divided by *M* from step 3).
[c]The peak of the linear fit pulse response *p*(k) from step 3).

92.8.3.1 Transmitter differential output return loss

The differential output return loss, in dB, of the transmitter shall meet Equation (92–1). This output impedance requirement applies to all valid output levels. The reference impedance for differential return loss measurements shall be 100 Ω.

$$Return\_loss(f) \geq \{ TBD \} \quad \text{(dB)} \qquad (92–1)$$

where

f	is the frequency in GHz
Return_loss(f)	is the return loss at frequency f

92.8.3.2 Transmitter noise parameter measurements

The far-end transmitter output noise is an additional source of noise to the cable assembly's integrated crosstalk noise (ICN) specified in 92.10.7. The far-end transmitter output noise parameter is characterized using two reference channels; a "low-loss" cable assembly with insertion loss on the reference pair of TBD dB ± TBD dB at 12.8906 GHz and cable assembly integrated crosstalk noise (ICN) meeting the requirements of 92.10.7 and a "high-loss" cable assembly with insertion loss on the reference pair of TBD dB ± TBD dB at 12.8906 GHz and cable assembly ICN meeting the requirements of 92.10.7. The far-end transmitter output noise is characterized as a deviation from the cable assembly ICN using the following procedure:

1) Compute the far-end integrated crosstalk noise σ_{fx} into the reference lane of the cable assembly using the methodology of 92.10.7 and the parameters in Table 92–11.
2) Denote σ_l as the far-end ICN for the low-loss cable assembly.
3) Denote σ_h as the far-end ICN for the high-loss cable assembly.
4) The transmitter under test is connected to one end of the reference cable assembly and the other end is connected to the cable assembly test fixture specified in 92.10.8.
5) All lanes of the cable assembly test fixture are terminated in the reference impedance with the reference lane connected to the measuring instrument.
6) The reference lane of the transmitter under test sends a square wave test pattern as specified in 83.5.10 while all other adjacent transmitter lanes send either scrambled idle or PRBS31.
7) A fixed point on the square wave test pattern is chosen and the RMS deviation from the mean voltage at this observation point is measured. The histogram for RMS noise measurement is 1 UI wide.
8) The measurement should not include the measurement system noise.

For the low-loss cable assembly, the measured RMS deviation from the cable assembly ICN due to the far-end transmitter output noise shall meet the values determined using Equation (92–2).

TBD (92–2)

For the high-loss cable assembly, the measured RMS deviation from the cable assembly ICN due to the far-end transmitter output noise shall meet the values determined using Equation (92–3).

TBD (92–3)

92.8.3.3 Transmitter output waveform

The 100GBASE-CR4 transmit function includes programmable equalization to compensate for the frequency-dependent loss of the channel and facilitate data recovery at the receiver. The functional model

for the transmit equalizer is the three tap transversal filter shown in Figure 92–3.

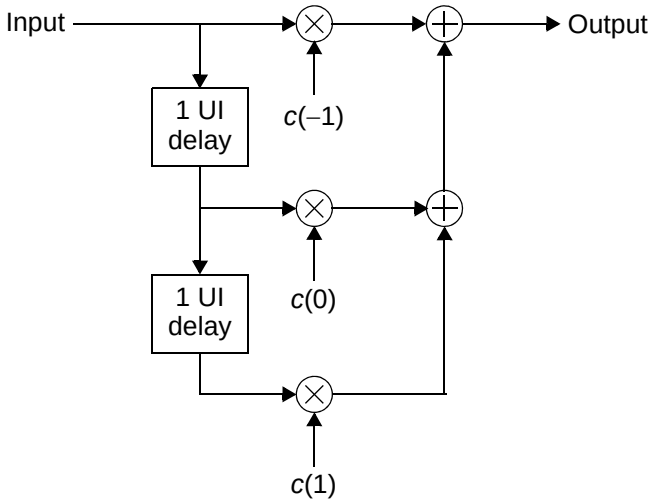


Figure 92–3—Transmit equalizer functional model

The state of the transmit equalizer and hence the transmitted output waveform may be manipulated via the PMD control function defined in 92.7.12 or via the management interface. The transmit function responds to a set of commands issued by the link partner's receive function and conveyed by a back-channel communications path.

This command set includes instructions to

a) Increment coefficient $c(i)$.
b) Decrement coefficient $c(i)$.
c) Hold coefficient $c(i)$ at its current value.
d) Set the coefficients to a pre-defined value (preset or initialize).

In response, the transmit function relays status information to the link partner's receive function. The status messages indicate that

a1) The requested update to coefficient $c(i)$ has completed (updated).
b1) Coefficient $c(i)$ is at its minimum value.
c1) Coefficient $c(i)$ is at its maximum value.
d1) Coefficient $c(i)$ is ready for the next update request (not_updated).

The requirements for the 100GBASE-CR4 transmit equalizer are intended to be similar to the requirements for 100GBASE-KR specified in 93.8.1.6. However, the signal path from the transmit function to TP2 introduces frequency-dependent loss and phase shift that distorts the signal and makes it difficult to accurately characterize equalizer performance at TP2 using the methodology specified for 10GBASE-KR. The following process enables accurate characterization of the equalizer performance at TP2 by determining and correcting for the frequency-dependent loss and phase shift of the signal path from the transmit function to TP2.

1) The transmitter under test is preset as specified in 72.6.10.2.3.1 such that $c(-1)$ and $c(1)$ are zero and $c(0)$ is its maximum value.
2) Capture at least one complete cycle of the test pattern PRBS9 as specified in 83.5.10 at TP2 per 92.8.3.3.4.

3) Compute the linear fit to the captured waveform and the linear fit pulse response $p(k)$ per 92.8.3.3.5.
4) Define t_x to be the time where the rising edge of the linear fit pulse, p, from step 3) crosses 50% of its peak amplitude.
5) Sample the linear fit pulse, p, at symbol-spaced intervals relative to the time $t_0 = t_x + 0.5$ UI, interpolating as necessary to yield the sampled pulse p_i.
6) Use p_i to compute the vector of coefficients, w, of a N_w-tap symbol-spaced transversal filter that equalizes for the transfer function from the transmit function to TP2 per 92.8.3.3.6.

The parameters of the pulse fit and the equalizing filter are given in Table 92–6. The Steady state voltage, the sum of linear fit pulse response, $p(k)$, from step 3) divided by M from step 3), shall be greater than TBD V and less than or equal to TBD V. The peak of the linear fit pulse response from step 3) shall be greater than TBD × Steady state voltage. The RMS value of the error between the linear fit from step 3) and the measured waveform, e, normalized to the peak value of the pulse, p, must be no greater than 0.037.

Table 92–6—Normalized transmit pulse template

Description	Symbol	Value	Units
Linear fit pulse length	N_p	TBD	UI
Linear fit pulse delay	D_p	TBD	UI
Equalizer length	N_w	TBD	UI
Equalizer delay	D_w	TBD	UI

For each configuration of the transmit equalizer

7) Configure the transmitter under test as required by the test.
8) Capture at least one complete cycle of the test pattern PRBS9 as specified in 83.5.10 at TP2 per 92.8.3.3.4.
9) Compute the linear fit to the captured waveform and the linear fit pulse response $p(k)$ per 92.8.3.3.5.
10) Define t_x to be the time where the rising edge of the linear fit pulse response, $p(k)$, from step 9 crosses 50% of its peak amplitude.
11) Sample the linear fit pulse response, $p(k)$, at symbol-spaced intervals relative to the time $t_0 = t_x + 0.5$ UI, interpolating as necessary to yield the sampled pulse p_i.
12) Equalize the sampled pulse p_i using the coefficient vector, w, computed in step 6) per 92.8.3.3.6 to yield the equalized pulse q_i.

The RMS value of the error between the linear fit from step 9) and the measured waveform, e, normalized to the peak value of the pulse, p, must be no greater than 0.037.

The normalized amplitude of coefficient $c(-1)$ is the value of q_i at time $t_0 + (D_p - 1)$ UI. The normalized amplitude of coefficient $c(0)$ is the value of q_i at time $t_0 + D_p$ UI. The normalized amplitude of coefficient $c(1)$ is the value of q_i at time $t_0 + (D_p + 1)$ UI.

92.8.3.3.1 Coefficient initialization

When the PMD enters the INITIALIZE state of the Training state diagram (Figure 72–5) or receives a valid request to “initialize” from the link partner, the coefficients of the transmit equalizer shall be configured such that the ratio $(c(0)+c(1)-c(-1))/(c(0)+c(1)+c(-1))$ is TBD ±10% and the ratio

$(c(0)–c(1)+c(–1))/(c(0)+c(1)+c(–1))$ is TBD ±–10%. These requirements apply upon the assertion a coefficient status report of "updated" for all coefficients.

92.8.3.3.2 Coefficient step size

The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to "increment" that coefficient shall be between TBD and TBD. The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to "decrement" that coefficient shall be between TBD and TBD.

The change in the normalized amplitude of the coefficient is defined to be the difference in the value measured prior to the assertion of the "increment" or "decrement" request (e.g., the coefficient update request for all coefficients is "hold") and the value upon the assertion of a coefficient status report of "updated" for that coefficient.

92.8.3.3.3 Coefficient range

When sufficient "increment" or "decrement" requests have been received for a given coefficient, the coefficient will reach a lower or upper bound based on the coefficient range or restrictions placed on the minimum Steady state differential output voltage or the maximum peak-to-peak differential output voltage.

With $c(–1)$ set to zero and both $c(0)$ and $c(1)$ having received sufficient "decrement" requests so that they are at their respective minimum values, the ratio $(c(0) – c(1))/(c(0) + c(1))$ shall be greater then or equal to TBD.

With $c(1)$ set to zero and both $c(–1)$ and $c(0)$ having received sufficient "decrement" requests so that they are at their respective minimum values, the ratio $(c(0) – c(–1))/(c(0) + c(–1))$ shall be greater then or equal to TBD.

Note that a coefficient may be set to zero by first asserting a coefficient preset request and then manipulating the other coefficients as required by the test.

92.8.3.3.4 Waveform acquisition

The transmitter under test repetitively transmits the specified test pattern. The waveform shall be captured with an effective sample rate that is M times the signaling rate of the transmitter under test. The value of M shall be an integer not less than TBD. Averaging multiple waveform captures is recommended.

The captured waveform shall represent an integer number of repetitions of the test pattern totaling N bits. Hence the length of the captured waveform should be MN samples. The waveform should be aligned such that the first M samples of waveform correspond to the first bit of the test pattern, the second M samples to the second bit, and so on.

92.8.3.3.5 Linear fit to the waveform measurement at TP2

Given the captured waveform $y(k)$ and corresponding aligned symbols $x(\mathrm{n})$ derived from the procedure defined in 92.8.3.3.4, define the M-by-N waveform matrix Y as shown in Equation (92–4).

$$Y = \begin{bmatrix} y(1) & y(M+1) & \dots & y(M(N-1)+1) \\ y(2) & y(M+2) & \dots & y(M(N-1)+2) \\ \dots & \dots & \dots & \dots \\ y(M) & y(2M) & \dots & y(MN) \end{bmatrix} \quad (92–4)$$

Rotate the symbols vector x by the specified pulse delay D_p to yield x_r as shown in Equation (92–5).

$$x_r = \left[x(D_p + 1)\ \ x(D_p + 2) \ldots\ x(N)\ x(1)\ \ldots\ x(D_p)\right] \tag{92–5}$$

Define the matrix X to be an N-by-N matrix derived from x_r as shown in Equation (92–6).

$$X = \begin{bmatrix} x_r(1) & x_r(2) & \ldots & x_r(N) \\ x_r(N) & x_r(1) & \ldots & x_r(N-1) \\ \ldots & \ldots & \ldots & \ldots \\ x_r(2) & x_r(3) & \ldots & x_r(1) \end{bmatrix} \tag{92–6}$$

Define the matrix X_1 to be the first N_p rows of X concatenated with a row vector of ones of length N. The M-by-$(N_p + 1)$ coefficient matrix, P, corresponding to the linear fit is then defined by Equation (92–7). The superscript "T" denotes the matrix transpose operator.

$$P = YX_1^T(X_1X_1^T)^{-1} \tag{92–7}$$

The error waveform, $e(k)$, is then read column-wise from the elements of E as shown in Equation (92–8).

$$E = PX_1 - Y = \begin{bmatrix} e(1) & e(M+1) & \ldots & e(M(N-1)+1) \\ e(2) & e(M+2) & \ldots & e(M(N-1)+2) \\ \ldots & \ldots & \ldots & \ldots \\ e(M) & e(2M) & \ldots & e(MN) \end{bmatrix} \tag{92–8}$$

Define P_1 to be a matrix consisting of the first N_p columns of the matrix P as shown in Equation (92–9). The linear fit pulse response, $p(k)$, is then read column-wise from the elements of P_1.

$$P_1 = \begin{bmatrix} p(1) & p(M+1) & \ldots & p(M(N_p-1)+1) \\ p(2) & p(M+2) & \ldots & p(M(N_p-1)+2) \\ \ldots & \ldots & \ldots & \ldots \\ p(M) & p(2M) & \ldots & p(MN_p) \end{bmatrix} \tag{92–9}$$

92.8.3.3.6 Transfer function between the transmit function and TP2

Rotate the sampled pulse response p_i by the specified equalizer delay D_w to yield p_r as shown in Equation (92–10).

$$p_r = \left[p_i(D_w + 1)\ \ p_i(D_p + 2) \ldots\ p_i(N_p)\ p_i(1)\ \ldots\ p_i(D_w)\right] \tag{92–10}$$

Define the matrix P_2 to be an N_p-by-N_p matrix derived from p_r as shown in Equation (92–11).

$$P_2 = \begin{bmatrix} p_r(1) & p_r(N_p) & \ldots & p_r(2) \\ p_r(2) & p_r(1) & \ldots & p_r(3) \\ \ldots & \ldots & \ldots & \ldots \\ p_r(N_p) & p_r(N_p-1) & \ldots & p_r(1) \end{bmatrix} \tag{92–11}$$

Define the matrix P_3 to be the first N_w columns of P_2. Define a unit pulse column vector x_p of length N_p. The value of element $x_p(D_p + 1)$ is 1 and all other elements have a value of 0. The vector of filter coefficients w that equalizes p_i is then defined by Equation (92–12).

$$w = (P_3^T P_3)^{-1} P_3^T x_p \quad (92–12)$$

Given the column vector of equalizer coefficients, w, the equalized pulse response q_i is determined by Equation (92–13).

$$q_i = P_3 w \quad (92–13)$$

92.8.3.4 Insertion loss TP0 to TP2 or TP3 to TP5

Transmitter measurements and tests defined in Table 92–5 are made at TP2 or TP3 using the test fixture of Figure 92–5, or its equivalent. The maximum insertion loss from TP0 to TP2 or TP3 to TP5 including the test fixture is determined using Equation (92–14). Note that in Annex 92A, the insertion loss from TP0 to TP2 or from TP3 to TP5 is 10 dB at 12.8906 GHz

$$Insertion\_loss(f) \leq \{TBD\} \quad \text{(dB)} \quad (92–14)$$

where

f	is the frequency in GHz
$Insertion\_loss(f)$	is the insertion loss at frequency f

The maximum insertion loss of TP0 to TP2 or TP3 to TP5 is illustrated in Figure 92–4.

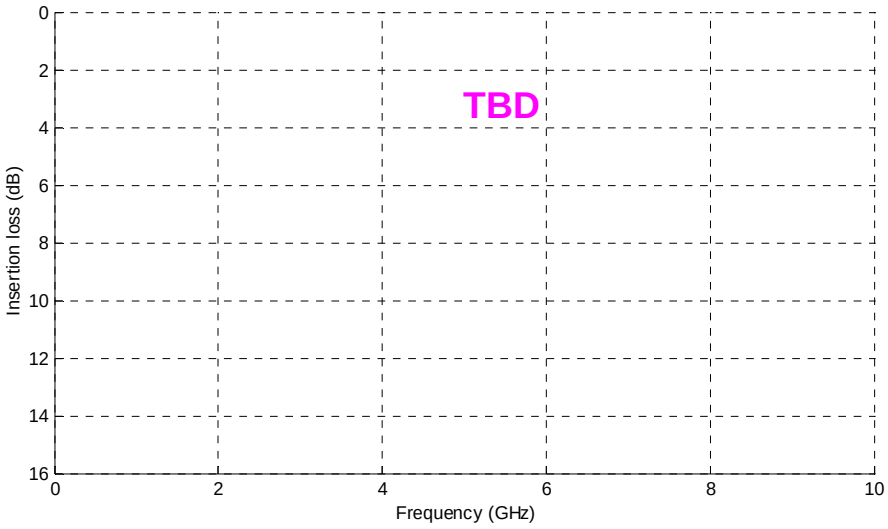


Figure 92–4—Maximum insertion loss TP0 to TP2 or TP3 to TP5

92.8.3.5 Test fixture

The test fixture of Figure 92–5, or its equivalent, is required for measuring the transmitter specifications in 92.8.3 at TP2 and the receiver return loss at TP3. TP2 and TP3 are illustrated in Figure 92–2. Figure 92–5 illustrates the test fixture attached to TP2 or TP3.

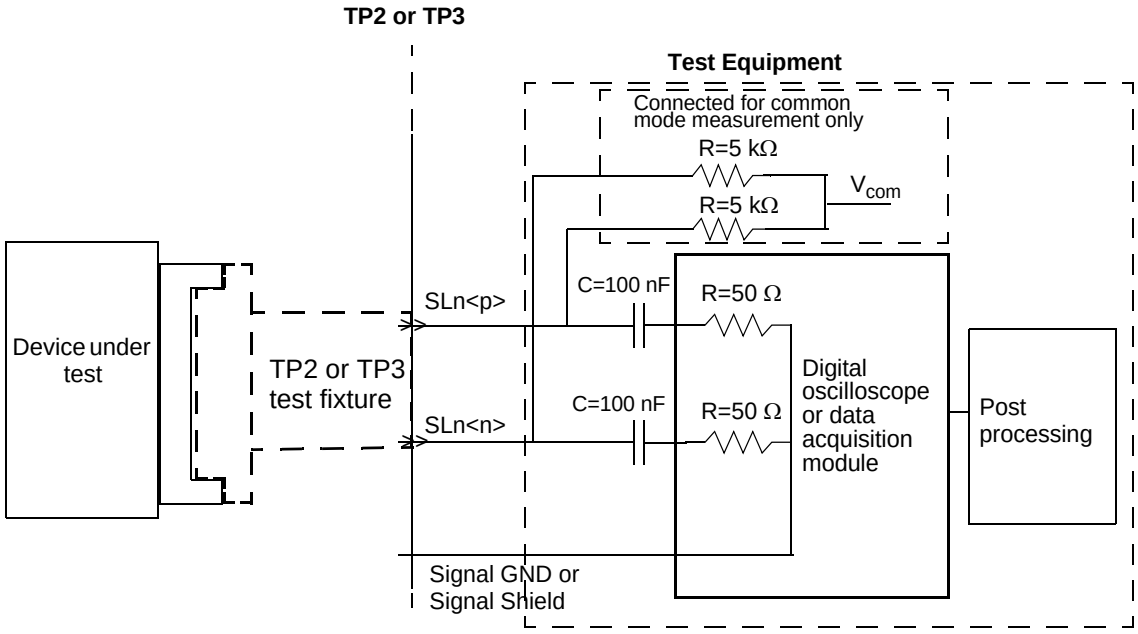


Figure 92–5—Transmitter and receiver test fixture

92.8.3.6 Test fixture impedance

The differential return loss, in dB, of the test fixture is specified in a mated state and shall meet the requirements of 92.10.9.2.

92.8.3.7 Test fixture insertion loss

The reference test fixture printed circuit board insertion loss is given in Equation (92–15) and shall be used. The effects of differences between the insertion loss of an actual test fixture and the reference insertion loss are to be accounted for in the measurements.

$$IL_{tfref}(f) = TBD \quad \text{(dB)} \qquad (92\text{–}15)$$

for TBD $\le f \le$ TBD GHz

where

f is the frequency in GHz
$IL_{tfref}(f)$ is the reference test fixture PCB insertion loss at frequency f

Editor's note (to be removed prior to final publication):

In Annex 92A, the insertion loss of the test fixture printed circuit board is 1.25 dB at 12.8906 GHz.

92.8.3.8 Transmitter output jitter

Even-odd jitter is measured from the two symbols in the middle of a sequence of no fewer than 8 symbols of alternating polarity. If PRBS9 is the test pattern, a suitable sequence may be found starting at either bit 161 or bit 383 where bits 1 to 9 are the run of 9 ones. A correct measurement of even-odd jitter requires that the period of the test pattern is an even number of symbols. When the base pattern period is an odd number of symbols, such as PRBS9, the test pattern for the purpose of even-odd jitter measurement must be two periods of the base pattern.

Even-odd jitter is defined to be half of the magnitude of the difference between the mean width of the positive pulse and the mean width of the negative pulse. The reference voltage for pulse width measurements is the mid-point between the positive pulse amplitude and the negative pulse amplitude. The pulse amplitude is defined to be the mean amplitude of the pulse within the central 20% of the nominal unit interval.

Even-odd jitter shall be less than or equal to 0.035 UI regardless of the transmit equalization setting.

NOTE—Even-odd jitter has been referred to as duty cycle distortion by other Physical Layer specifications for operation over electrical backplane or twinaxial copper cable assemblies (see 72.7.1.9). The term even-odd jitter is introduced to distinguish it from the duty cycle distortion referred to by Physical Layer specifications for operation over fiber optic cabling.

The total jitter (TJ) of a signal is defined as the range (the difference between the lowest and highest values) of sampling times around the signal transitions for which the BER at these sampling times is greater than or equal to 10^{-12}. Data dependent jitter (DDJ) is characterized using the procedure defined in 85.8.3.8.

The difference between TJ and DDJ shall be less than or equal to 0.28 UI regardless of the transmit equalization setting.

Editor's note (to be removed prior to final publication):

Clause 85 was not clear on what was meant by "excluding" DDJ from TJ. In this draft, it has been interpreted to be the value of TJ as defined above minus the value of DDJ as defined in 85.8.3.8.

The effective random jitter (RJ) of a signal is defined to be the difference between the TJ and effective deterministic jitter (DJ). Effective DJ is derived from a fit of the measured jitter distribution to a dual-Dirac mathematical model. The fit is computed as follows.

a) Measure the jitter J_n which is defined to be the range of sampling times around the signal transitions for which the BER at these sampling times is greater than or equal to BER_n. Measure two values J_0 and J_1 where BER_0 is less than BER_1 e.g. 10^{-9} and 10^{-5}.
b) For each BER_n, determine the associated Q_n from the inverse normal cumulative probability distribution adjusted for an assumed transition density of 0.5. For example, Q_n is 5.77 if BER_n is 10^{-9} and Q_n is 3.94 if BER_n is 10^{-5}.
c) Calculate the effective DJ as $(Q_0 J_1 - Q_1 J_0) / (Q_0 - Q_1)$.

The effective RJ shall be less than or equal to 0.15 UI regardless of the transmit equalization setting.

Editor's note (to be removed prior to final publication):

Clause 72 specifies that transmit equalization is off (preset) for the measurement of jitter. Clause 85 does not specify the conditions for jitter measurement and this is interpreted to mean that it applies for all transmit equalization settings. There has been some discussion as to which approach is appropriate.

The effect of a single-pole high-pass filter with a 3 dB frequency of 10 MHz is applied to the jitter. The test pattern for TJ and RJ measurements is either PRBS31 (see 83.5.10) or scrambled idle (see 82.2.10). The voltage threshold for the measurement of BER or crossing times is the mid-point (0 V) of the AC-coupled differential signal.

92.8.3.9 Signaling rate range

The 100GBASE-CR4 MDI signaling rate shall be 25.78125 GBd ±100 ppm per lane. The corresponding unit interval is nominally 38.787879 ps.

92.8.4 Receiver characteristics at TP3 summary

The receiver characteristics shall meet the specifications summarized in Table 92–7 at TP3 unless otherwise noted.

Table 92–7—Receiver characteristics at TP3 summary

Parameter	Subclause reference	Value	Units
Bit error ratio	92.8.4.3	10^{-12} or better	
Signaling rate, per lane	92.8.4.4	25.78125 ± 100 ppm	GBd
Unit interval (UI) nominal	92.8.4.4	38.787879	ps
Differential peak-to-peak input amplitude tolerance (max)	72.7.2.4	1200	mV
Differential input return loss (min)[a]	92.8.4.1	Equation (92–16)	dB
Differential to common-mode input return loss		10 min from 0.01 GHz to 25 GHz	dB

[a]Relative to 100 Ω differential.

92.8.4.1 Receiver differential input return loss

The differential input return loss, in dB, of the receiver shall meet Equation (92–16). This return loss requirement applies at all valid input levels. The reference impedance for differential return loss measurements is 100 Ω.

$$Return\_loss(f) \geq \begin{cases} 12 - 1.24\sqrt{f} & 0.01 \leq f < 10.31 \\ 6.3 - 13\log_{10}(f/13.75) & 10.31 \leq f \leq 25 \end{cases} \quad (\text{dB}) \qquad (92\text{–}16)$$

where

f is the frequency in GHz
$Return\_loss(f)$ is the return loss at frequency f

92.8.4.2 Receiver interference tolerance test

The receiver interference tolerance of each lane shall comply with both test 1 and test 2 using the parameters of Table 92–8 when measured according to the requirements of 92.8.4.2.1 to 92.8.4.2.5.

Table 92–8—100GBASE-CR4 interference tolerance parameters

Parameter	Test 1 values	Test 2 values	Units
Maximum BER	10^{-12}	10^{-12}	
Fitted insertion loss coefficients	a_1 = TBD a_2 = TBD a_4 = TBD	a_1 = TBD a_2 = TBD a_4 = TBD	dB/√GHz dB/GHz dB/GHz^2
Applied SJ[a] (peak-to-peak)	TBD	TBD	UI
Applied RJ[b] (peak-to-peak)	TBD	TBD	UI
Applied DCD (peak-to-peak)	TBD	TBD	UI
Calibrated far-end crosstalk (RMS)	TBD	TBD	mV
Calibrated ICN (RMS) – σ_{nx} [c]	TBD	TBD	mV

[a]Applied SJ frequency >15 MHz, specified at TP0.
[b]Applied random jitter at TP0 is specified at 10^{-12}.
[c]σ_{nx} is subtracted from Calibrated ICN (RMS) in square root of the sum of the squares sense.

92.8.4.2.1 Test setup

The interference tolerance test is performed with the setup shown in Figure 92–6. The requirements of this subclause are verified at the pattern generator connection (PGC) or test references in Figure 92–6 and Figure 92–7. The lanes under test (LUT) are illustrated in Figure 92–6 and Figure 92–7. The cable assembly receive lanes are terminated in 100 Ω differentially.

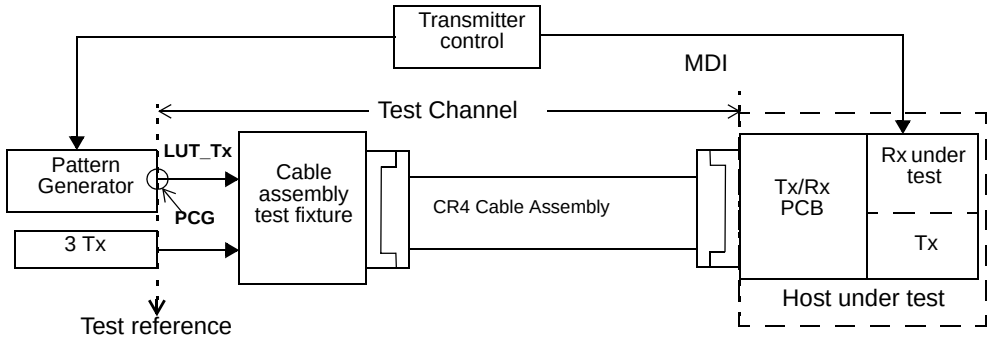


Figure 92–6—Interference tolerance test setup

92.8.4.2.2 Test channel

The test channel consists of the following:

a) A cable assembly
b) A cable assembly test fixture
c) A connecting path from the pattern generator to the cable assembly test fixture

92.8.4.2.3 Test channel calibration

The insertion loss, near-end integrated crosstalk noise, and far-end crosstalk of the test channels are characterized at the test references as illustrated in Figure 92–6 using the cable assembly test fixtures specified in 92.10.8.

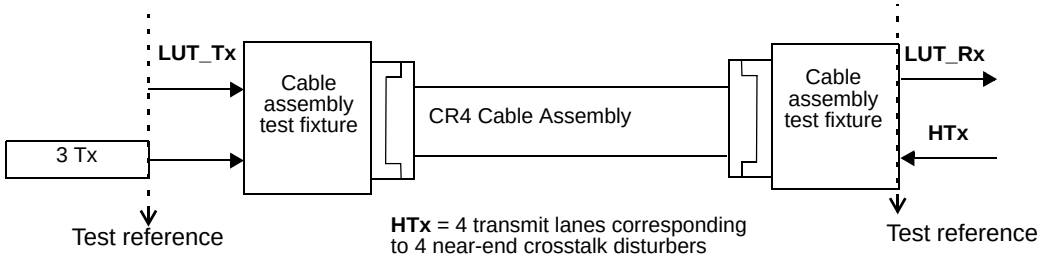


Figure 92–7—Test channel calibration

The fitted insertion loss coefficients of the lane under test (LUT), derived using the fitting procedure in 92.10.2, shall meet the test values in Table 92–8. It is recommended that the deviation between the insertion loss and the fitted insertion loss be as small as practical and that the fitting parameters be as close as practical to the values given in Table 92–8.

The MDNEXT is measured from points HTx to adjacent point LUT_Rx in Figure 92–7. HTx is the set of 4 transmit lanes of the device under test corresponding to the 4 near-end crosstalk disturbers with the parameters given in Table 92–11. The RMS value of the integrated MDNEXT crosstalk noise, determined using Equation (92–29) through Equation (92–33), shall meet the test values in Table 92–8.

The far-end crosstalk disturbers consist of 100GBASE-CR4 transmitters. It is recommended that the transition time, equalization setting, and path from the far-end crosstalk disturbers to the cable assembly test fixture emulate the pattern generator as much as practical. For 100GBASE-CR4 test channels, the crosstalk that is coupled into a receive lane is from three transmitters. The disturber transmitters send either scrambled idle codes or PRBS31. The amplitudes of each of the disturbers should not deviate more than 3 dB from the mean of the disturber amplitudes. The amplitudes of the disturbers should be such that the calibrated far-end crosstalk in Table 92–8 is met in the calibration setup at the LUT point with no signal applied at the PGC, and HTx and PGC terminated in 100 Ω differentially.

92.8.4.2.4 Pattern generator

The pattern generator transmits data to the device under test. Its output amplitude shall be no more than TBD mV peak-to-peak differential when measured on an alternating one zero pattern. The rise and fall times of the pattern generator, as defined in 72.7.1.7, are TBD ps. If the rise and fall times of the pattern generator, T_r, are less than TBD ps the value of a_4 in Table 92–8 is increased by da_4 from Equation (92–17).

TBD (92–17)

where T_r is the rise time in ps.

The pattern generator shall meet the jitter specification in Table 92–8. The output waveform of the pattern generator shall comply to 93.8.1.

92.8.4.2.5 Test procedure

For 100GBASE-CR4 testing, the pattern generator is first configured to transmit the training pattern defined in 72.6.10.2. During this initialization period, the device under test (DUT) configures the pattern generator equalizer, via transmitter control, to the coefficient settings it would select using the protocol described in 72.6.10 and the receiver will be tuned using its optimization method.

After the pattern generator equalizer has been configured and the receiver tuned, the pattern generator is set to transmit test pattern 3 as defined in 86.8.2. The receiver under test shall meet the target BER listed in Table 92–8. During the tests, the disturbers transmit at their calibrated level and all of the transmitters in the device under test transmit either scrambled idle characters or PRBS31, with the maximum compliant amplitude and equalization turned off (preset condition).

92.8.4.3 Bit error ratio

The receiver shall operate with a BER 10^{-12} or better when receiving a compliant transmit signal, as defined in 92.8.3, through a compliant cable assembly as defined in 92.10.

92.8.4.4 Signaling rate range

A 100GBASE-CR4 receiver shall comply with the requirements of 92.8.4.3 for any signaling rate in the range 25.78125 GBd ± 100 ppm. The corresponding unit interval is nominally 38.787879 ps.

92.8.4.5 AC coupling

The 100GBASE-CR4 receivers are AC coupled. AC coupling shall be part of the receive function for Style-2 100GBASE-CR4 connectors. For Style-1 100GBASE-CR4 plug connectors, the receive lanes are AC coupled; the coupling capacitors shall be within the plug connectors. It should be noted that there may be various methods for AC coupling in actual implementations. The low frequency 3 dB cutoff of the AC coupling shall be less than TBD kHz.

It is recommended that the value of the coupling capacitors be 100 nF. This will limit the inrush currents and baseline wander.

92.9 Channel characteristics

The 100GBASE-CR4 channel is defined between TP0 and TP5 to include the transmitter and receiver differential controlled impedance printed circuit board and the cable assembly as illustrated in Figure 92–2. The channel parameters insertion loss, insertion loss deviation (ILD), insertion loss to crosstalk ratio, and the transmitter and receiver differential controlled impedance printed circuit boards for each differential lane are provided informatively in 92A.4 through 92A.8.

92.10 Cable assembly characteristics

The 100GBASE-CR4 cable assembly contains insulated conductors terminated in a connector at each end for use as a link segment between MDIs. This cable assembly is primarily intended as a point-to-point interface of up to 5 m between network ports using controlled impedance cables. All cable assembly measurements are to be made between TP1 and TP4 with cable assembly test fixtures as specified in 92.10.8 and illustrated in Figure 92–12. These cable assembly specifications are based upon twinaxial cable characteristics, but other cable types are acceptable if the specifications of 92.10 are met.

Table 92–9 provides a summary of the cable assembly differential characteristics at 12.8906 GHz and references to the subclauses addressing each parameter.

Table 92–9—Cable assembly differential characteristics summary

Description	Reference	Value	Unit
Maximum insertion loss at 12.8906 GHz	92.10.2	22.64	dB
Minimum insertion loss at 12.8906 GHz		4	dB
Insertion loss deviation at 12.8906 GHz	92.10.3	max = 2.97 min = –2.97	dB
Minimum return loss at 12.8906 GHz	92.10.5	6.0	dB
MDNEXT loss	92.10.5	Equation (92–25)	dB
MDFEXT loss	92.10.6	Equation (92–26)	dB
Maximum integrated crosstalk noise	92.10.7	Equation (92–32)	mV

92.10.1 Characteristic impedance and reference impedance

The nominal differential characteristic impedance of the cable assembly is 100 Ω. The differential reference impedance for cable assembly specifications shall be 100 Ω.

92.10.2 Cable assembly insertion loss

The fitted cable assembly insertion loss $IL_{fitted}(f)$ as a function of frequency f is defined in Equation (92–18).

$$IL_{\text{fitted}}(f) = a_1\sqrt{f} + a_2 f + a_4 f^2 \quad \text{(dB)} \tag{92–18}$$

where

f is the frequency in GHz
$IL_{\text{fitted}}(f)$ is the fitted cable assembly insertion loss at frequency f

Given the cable assembly insertion loss measured between TP1 and TP4 is at N uniformly-spaced frequencies f_n spanning the frequency range 50 MHz to 18750 MHz with a maximum frequency spacing of 10 MHz, the coefficients of the fitted insertion loss are determined using Equation (92–19) and Equation (92–20).

Define the frequency matrix F as shown in Equation (92–19).

$$F = \begin{bmatrix} \sqrt{f_1} & f_1 & f_1^2 \\ \sqrt{f_2} & f_2 & f_2^2 \\ \cdots & \cdots & \cdots \\ \sqrt{f_N} & f_N & f_N^2 \end{bmatrix} \tag{92–19}$$

The polynomial coefficients a_1, a_2, and a_4 are determined using Equation (92–20). In Equation (92–20), T denotes the matrix transpose operator and IL is a column vector of the measured insertion loss values, IL_n at each frequency f_n.

$$\begin{bmatrix} a_1 \\ a_2 \\ a_4 \end{bmatrix} = (F^T F)^{-1} F^T IL \qquad (92–20)$$

The maximum allowed values of the polynomial coefficients a_1, a_2, and a_4 of the fitted cable assembly insertion loss of each pair of the 100GBASE-CR4 in Equation (92–18) and the maximum insertion loss at 12.8906 GHz shall meet the specifications summarized in Table 92–10 unless otherwise noted. The fitted insertion loss corresponding to one example of the maximum insertion loss at 12.8906 GHz and the maximum allowed values of a_1, a_2, and a_4 are illustrated in Figure 92–8.

Table 92–10—Maximum cable assembly insertion loss characteristics

Description	Value	Unit
Maximum insertion loss at 12.8906 GHz	22.64[a]	dB
Maximum fitted insertion loss coefficient a_1	TBD	dB/√GHz
Maximum fitted insertion loss coefficient a_2	TBD	dB/GHz
Maximum fitted insertion loss coefficient a_4	TBD	dB/GHz^2

[a]The limit on the maximum insertion loss at 12.8906 GHz precludes the coefficients a_1, a_2, and a_4 from simultaneous maximum values.

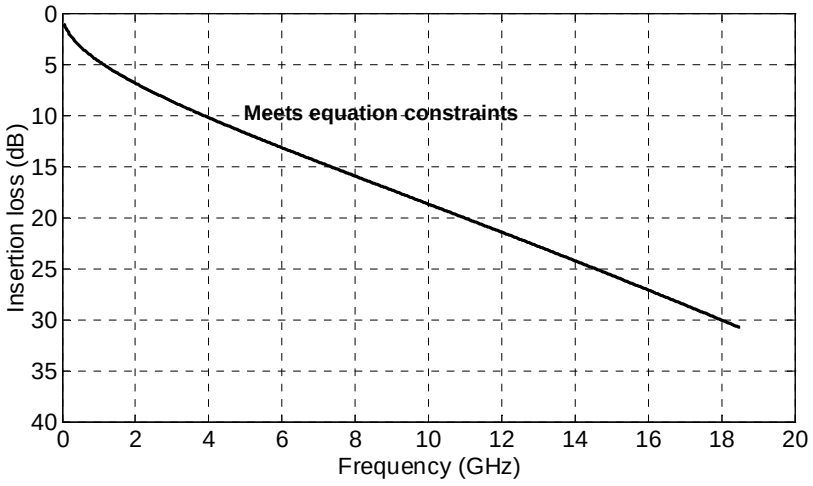


Figure 92–8—Example maximum cable assembly insertion loss

92.10.3 Cable assembly insertion loss deviation (ILD)

The cable assembly insertion loss deviation is the difference between the cable assembly insertion loss and the fitted cable assembly insertion loss determined using Equation (92–21).

$$ILD(f) = IL(f) - IL_{\text{fitted}}(f) \quad (92–21)$$

where

f is the frequency in GHz
ILD(*f*) is the cable assembly insertion loss deviation at frequency *f*

The *ILD* shall be within the region defined by Equation (92–22) and Equation (92–23). This includes the insertion loss of the differential cabling pairs and the cable assembly connectors.

$$ILD(f) \geq ILD_{\text{min}}(f) = -0.7 - 0.176f \quad \text{(dB)} \quad (92–22)$$

$$ILD(f) \leq ILD_{\text{max}}(f) = 0.7 + 0.176f \quad \text{(dB)} \quad (92–23)$$

for 0.05 GHz ≤ *f* ≤ 18.75 GHz

where

f is the frequency in GHz

The insertion loss deviation limits are illustrated in Figure 92–9.

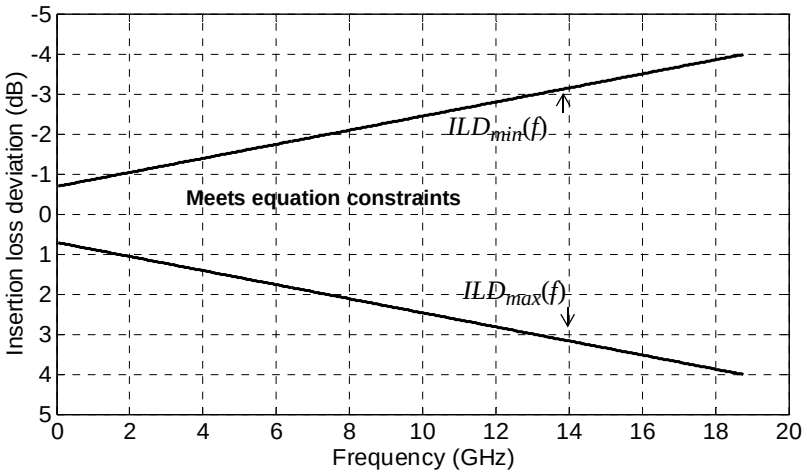


Figure 92–9—Maximum cable assembly insertion loss deviation

92.10.4 Cable assembly return loss

The return loss of each pair of the 100GBASE-CR4 cable assembly shall meet the values determined using Equation (92–24).

$$Return\_loss(f) \geq \begin{cases} 16.20 - 2\sqrt{f} & 0.05 \leq f < 4.1 \\ 10.59 - 13\log_{10}(f/5.5) & 4.1 \leq f \leq 25 \end{cases} \text{ (dB)} \quad (92\text{–}24)$$

where

f is the frequency in GHz
Return_loss(*f*) is the return loss at frequency *f*

The minimum cable assembly return loss is illustrated in Figure 92–10.

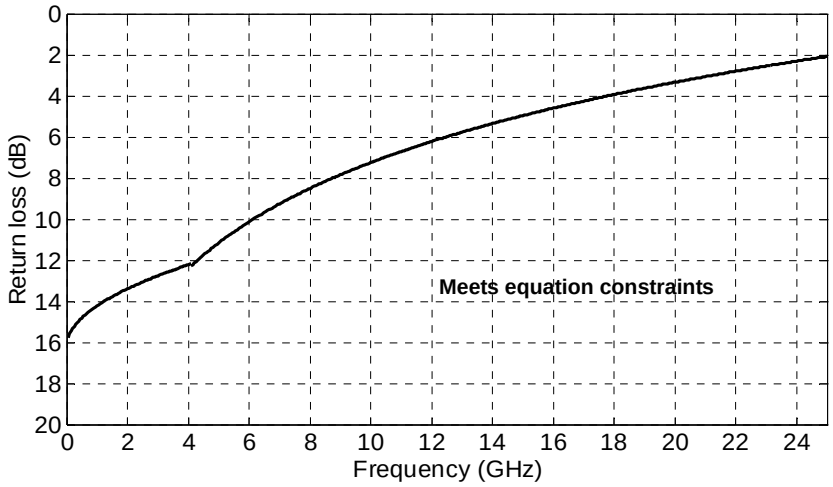


Figure 92–10—Minimum cable assembly return loss

92.10.5 Cable assembly multiple disturber near-end crosstalk (MDNEXT) loss

Since four lanes are used to transfer data between PMDs, the NEXT that is coupled into a receive lane will be from the four transmit lanes. Multiple Disturber Near-End Crosstalk (MDNEXT) loss is determined using the individual NEXT losses.

MDNEXT loss is determined from the four individual pair-to-pair differential NEXT loss values using Equation (92–25).

$$MDNEXT\_loss(f) = -10\log_{10}\left(\sum_{i=0}^{i=3} 10^{-NLi(f)/10}\right) \text{ (dB)} \quad (92\text{–}25)$$

for 0.05 GHz ≤ *f* ≤ 20 GHz

where

MDNEXT_loss(*f*) is the MDNEXT loss at frequency *f*
$NL_i(f)$ is the NEXT loss at frequency *f* of pair combination *i*, in dB

f	is the frequency in GHz
i	is the 0 to 3 (pair-to-pair combination)

92.10.6 Cable assembly multiple disturber far-end crosstalk (MDFEXT) loss

Since four lanes are used to transfer data between PMDs, the FEXT that is coupled into a data carrying lane will be from the three other lanes or nine other lanes in the same direction. MDFEXT loss is specified using the individual FEXT losses. MDFEXT loss is determined from the three individual pair-to-pair differential FEXT loss values using Equation (92–26).

$$MDFEXT\_loss(f) = -10\log_{10}\left(\sum_{i=0}^{i=2} 10^{-NLi(f)/10}\right) \text{ (dB)} \quad (92\text{–}26)$$

for 0.05 GHz ≤ f ≤ 20 GHz

where

$MDFEXT\_loss(f)$	is the MDFEXT loss at frequency f
$NL_i(f)$	is the FEXT loss at frequency f of pair combination i, in dB
f	is the frequency in GHz
i	is the 0 to 2 (pair-to-pair combination)

92.10.7 Cable assembly integrated crosstalk noise (ICN)

In order to limit multiple disturber crosstalk noise at a receiver, the cable assembly integrated crosstalk noise (ICN) is specified in relationship to the measured insertion loss. ICN is calculated from the MDFEXT and MDNEXT. Given the multiple disturber near-end crosstalk loss $MDNEXT\_loss(f)$ and multiple disturber far-end crosstalk loss $MDFEXT\_loss(f)$ measured over N uniformly-spaced frequencies f_n spanning the frequency range 50 MHz to 20000 MHz with a maximum frequency spacing of 10 MHz, the RMS value of the integrated crosstalk noise shall be determined using Equation (92–27) through Equation (92–31). The RMS crosstalk noise is characterized at the output of a specified receive filter utilizing a specified transmitter waveform and the measured multiple disturber crosstalk transfer functions. The transmitter and receiver filters are defined in Equation (92–27) and Equation (92–28) as weighting functions to the multiple disturber crosstalk in Equation (92–29) and Equation (92–30). The sinc function is defined by $\text{sinc}(x) = \sin(\pi x)/(\pi x)$.

Define the weight at each frequency f_n using Equation (92–27) and Equation (92–28).

$$W_{nt}(f_n) = (A_{nt}^2/f_b)\text{sinc}^2(f_n/f_b)\left[\frac{1}{1+(f_n/f_{nt})^4}\right]\left[\frac{1}{1+(f_n/f_r)^8}\right] \quad (92\text{–}27)$$

$$W_{ft}(f_n) = (A_{ft}^2/f_b)\text{sinc}^2(f_n/f_b)\left[\frac{1}{1+(f_n/f_{ft})^4}\right]\left[\frac{1}{1+(f_n/f_r)^8}\right] \quad (92\text{–}28)$$

where the equation parameters are given in Table 92–11.

Note that the 3 dB transmit filter bandwidths f_{nt} and f_{ft} are inversely proportional to the 20% to 80% rise and fall times T_{nt} and T_{ft} respectively. The constant of proportionality is 0.2365 (e.g., $T_{nt} f_{nt} = 0.2365$; with f_{nt} in hertz and T_{nt} in seconds). In addition, f_r is the 3 dB reference receiver bandwidth, which is set to 18.75 GHz.

The near-end integrated crosstalk noise σ_{nx} is calculated using Equation (92–29).

$$\sigma_{nx} = \left[2\Delta f \sum_{n} W_{nt}(f_n) 10^{-MDNEXT_{loss}(f_n)/10}\right]^{1/2} \quad (92\text{–}29)$$

The far-end integrated crosstalk noise σ_{fx} is calculated using Equation (92–30).

$$\sigma_{fx} = \left[2\Delta f \sum_{n} W_{ft}(f_n) 10^{-MDFEXT_{loss}(f_n)/10}\right]^{1/2} \quad (92\text{–}30)$$

where Δf is the uniform frequency step of f_n.

The total integrated crosstalk noise σ_x is calculated using Equation (92–31).

$$\sigma_x = \sqrt{\sigma_{nx}^2 + \sigma_{fx}^2} \quad (92\text{–}31)$$

The total integrated crosstalk noise for the cable assembly shall be computed using the parameters shown in Table 92–11.

Table 92–11—Cable assembly integrated crosstalk parameters

Description	Symbol	Value	Units
Symbol rate	f_b	25.78125	GBd
Near-end disturber peak differential output amplitude	A_{nt}	600	mV
Far-end disturber peak differential output amplitude	A_{ft}	600	mV
Near-end disturber 20% to 80% rise and fall times	T_{nt}	9.6	ps
Far-end disturber 20% to 80% rise and fall times	T_{ft}	9.6	ps

The total integrated crosstalk RMS noise voltage shall meet the values determined by Equation (92–32) illustrated in Figure 92–11.

$$\sigma_{x,\,ca} \leq \{ TBD \} \quad \text{(mV)} \quad (92\text{–}32)$$

where *IL* is the value of the cable assembly insertion loss in dB at 12.8906 GHz.

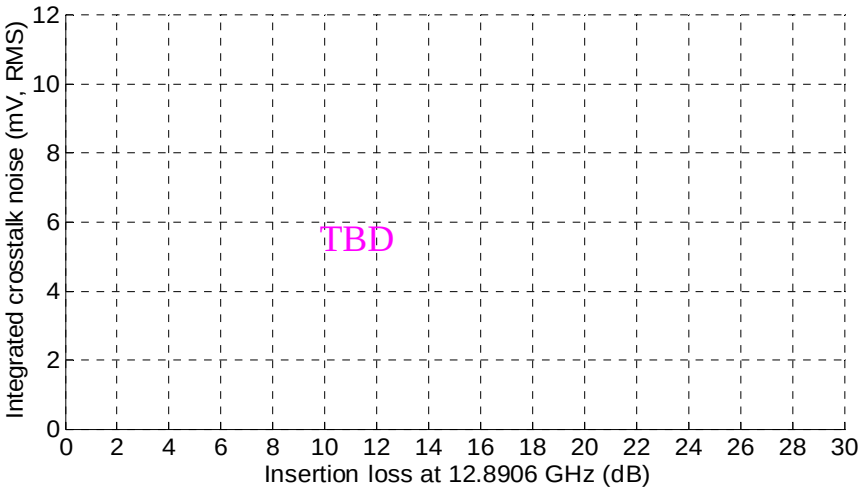


Figure 92–11—Integrated crosstalk noise limits

92.10.8 Cable assembly test fixture

The test fixture of Figure 92–12 or its equivalent, is required for measuring the cable assembly specifications in 92.10 at TP1 and TP4. TP1 and TP4 are illustrated in Figure 92–2 and Figure 92–12. The test fixture return loss is equivalent to the test fixture return loss specified in 92.8.3.6. The reference test fixture printed circuit board insertion loss is given in Equation (92–33). The effects of differences between the insertion loss of an actual test fixture and the reference insertion loss are to be accounted for in the measurements.

$$IL_{catf}(f) = TBD \quad \text{(dB)} \qquad (92–33)$$

for TBD GHz $\le f \le$ TBD GHz

Editor's note (to be removed prior to final publication):

In Annex 92A, the insertion loss of the cable assembly test fixture printed circuit board is 1.25 dB at 12.8906 GHz.

where

f is the frequency in GHz

$IL_{catf}(f)$ is the reference test fixture printed circuit board insertion loss at frequency f

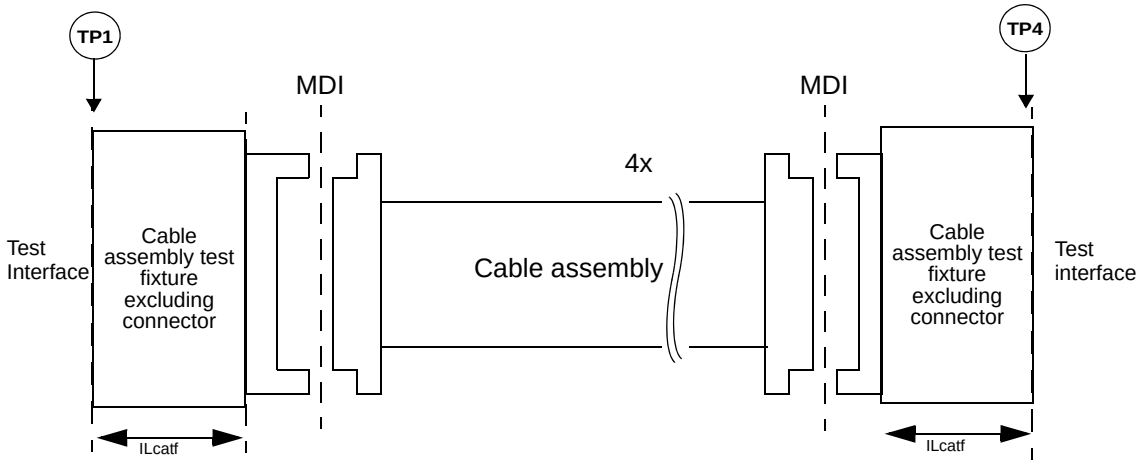


Figure 92–12—Cable assembly test fixtures

92.10.9 Mated test fixtures

The test fixtures of Figure 92–5 and Figure 92–12 are specified in a mated state illustrated in Figure 92–13.

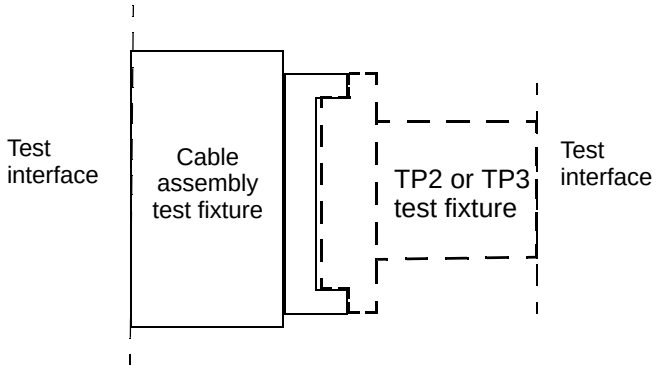


Figure 92–13—Mated test fixtures

92.10.9.1 Mated test fixtures insertion loss

The insertion loss of the mated test fixtures shall meet the values determined using Equation (92–34) and Equation (92–35).

$$IL(f) \geq IL_{MTFmin}(f) = TBD \quad TBD \leq f \leq TBD \quad \text{(dB)} \tag{92–34}$$

$$IL(f) \leq IL_{MTFmax}(f) = \{TBD\} \quad \text{(dB)} \tag{92–35}$$

where

f is the frequency in GHz
$IL(f)$ is the mated test fixture insertion loss at frequency f

Editor's note (to be removed prior to final publication):

In Annex 92A, the insertion loss of the mated test fixtures is 3.82 dB at 12.8906 GHz.

The mated test fixtures insertion loss limits are illustrated in Figure 92–14.

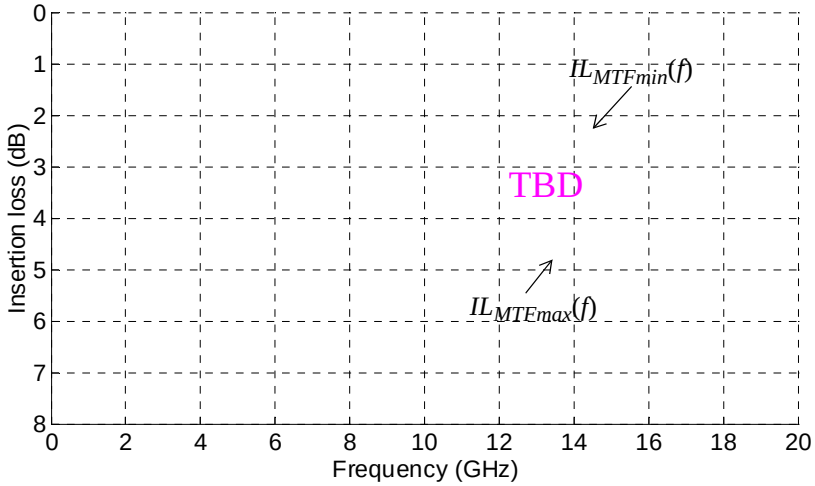


Figure 92–14—Mated test fixtures Insertion loss

92.10.9.2 Mated test fixtures return loss

The return loss of the mated test fixtures measured at either test fixture test interface shall meet the values determined using Equation (92–36).

$$Return\_loss(f) \geq \{TBD\} \quad \text{(dB)} \qquad (92–36)$$

where

f	is the frequency in GHz
$Return\_loss(f)$	is the return loss at frequency f

The mated test fixtures return loss is illustrated in Figure 92–15.

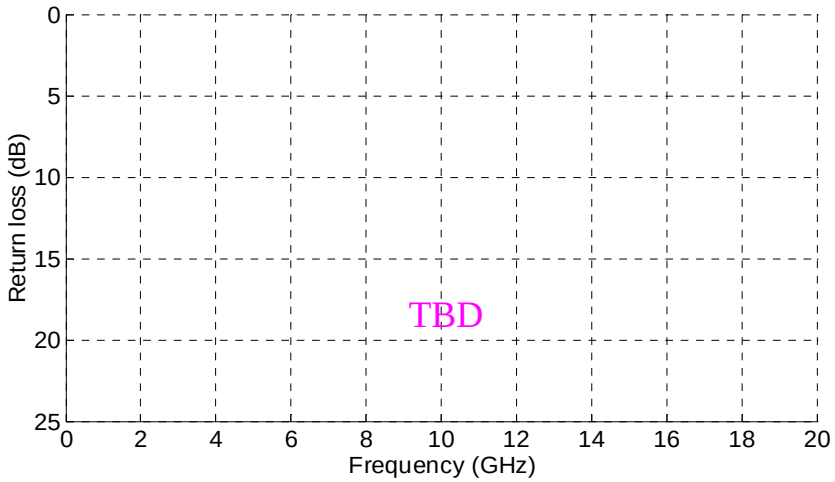


Figure 92–15—Mated test fixtures return loss

92.10.9.3 Mated test fixtures common-mode conversion loss

The common-mode conversion loss of the mated test fixtures measured at either test fixture test interface shall meet the values determined using Equation (92–37).

$$Conversion\_loss(f) \geq \{TBD\} \quad \text{(dB)} \qquad (92–37)$$

where

f is the frequency in GHz
Conversion_loss(*f*) is the return loss at frequency *f*

The mated test fixtures common-mode conversion loss is illustrated in Figure 92–16.

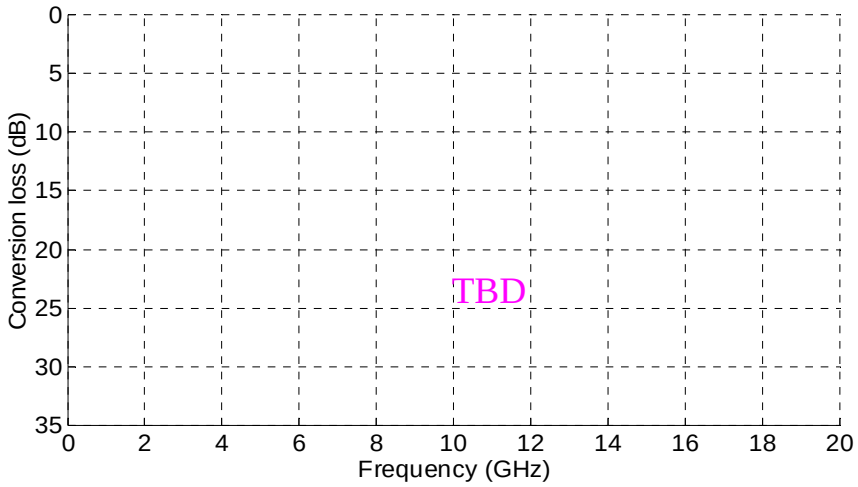


Figure 92–16—C

92.10.9.4 Mated test fixtures integrated crosstalk noise

The mated test fixtures integrated crosstalk RMS noise voltages for the single-disturber near-end crosstalk loss and the single-disturber far-end crosstalk loss are determined using Equation (92–27) through Equation (92–31) by substituting the single disturber near-end for the multiple disturber near-end crosstalk loss and the single disturber far-end crosstalk loss for the multiple disturber far-end crosstalk loss. The values of the mated test fixtures integrated crosstalk RMS noise voltages determined using Equation (92–27) through Equation (92–31) for the single-disturber near-end crosstalk loss, the single-disturber far-end crosstalk loss, the multiple disturber near-end crosstalk loss, and the multiple disturber far-end crosstalk loss shall meet the specifications in Table 92–12.

Table 92–12—Mated test fixtures integrated crosstalk noise

Parameter	100GBASE-CR4	Units
Near-end integrated crosstalk noise voltage (max, RMS)	TBD	mV
Far-end integrated crosstalk noise voltage (max, RMS)	TBD	mV
MDNEXT integrated crosstalk noise voltage (max, RMS)	TBD	mV
MDFEXT integrated crosstalk noise voltage (max, RMS)	TBD	mV

92.11 MDI specification

Editor's note (to be removed prior to final publication):

The following text is based on tracy_01_0312 per the motion passed at the March 2012 Task Force meeting.

"Move that QSFP28, per tracy_01_0312.pdf, be adopted as a baseline proposal for a type of MDI for 100GBASE-CR4. (M: N. Tracy, S: J. Goergen, Y: 55, N: 0, A: 2)"

Connectors meeting the requirements below shall be used as the mechanical interface between the PMD and the cable assembly. The plug connector shall be used on the cable assembly and the receptacle on the PHY.

The plug connector for each end of the cable assembly shall be the quad small form factor pluggable (QSFP28) with the mechanical mating interface defined by SFF-TBD and illustrated in Figure 92–17. The MDI connector shall be the quad small form factor pluggable (QSFP28) receptacle with the mechanical mating interface defined by SFF-TBD and illustrated in Figure 92–18. These connectors have contact assignments matching that in Table 92–13.

The plug connectors on the receive lanes are AC coupled, i.e. the coupling capacitors are contained within the plug connectors..

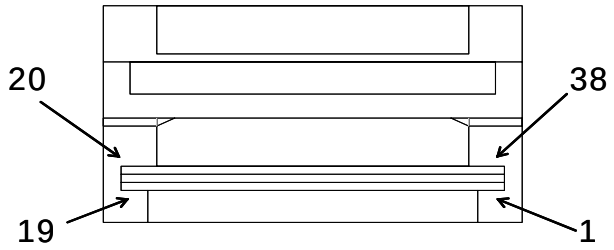


Figure 92–17—Example cable assembly plug

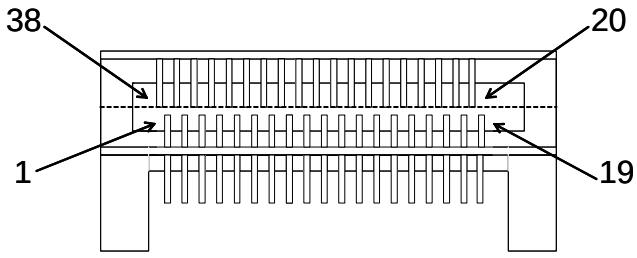


Figure 92–18—Example MDI board receptacle

Table 92–13—100GBASE-CR4 lane to MDI connector contact mapping

Tx lane	MDI connector contact	Rx lane	MDI connector contact
signal gnd	S1	signal gnd	S13
SL1<n>	S2	DL2<p>	S14
SL1<p>	S3	DL2<n>	S15
signal gnd	S4	signal gnd	S16
SL3<n>	S5	DL0<p>	S17
SL3<p>	S6	DL0<n>	S18
signal gnd	S7	signal gnd	S19
SL2<p>	S33	DL1<n>	S21
SL2<n>	S34	DL1<p>	S22
signal gnd	S35	signal gnd	S23
SL0<p>	S36	DL3<n>	S24
SL0<n>	S37	DL3<p>	S25
signal gnd	S38	signal gnd	S26

92.12 Environmental specifications

All equipment subject to this clause shall conform to the applicable requirements of 14.7.

92.13 Protocol implementation conformance statement (PICS) proforma for Clause 92, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-CR4[10]

92.13.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 92, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-CR4, shall complete the following protocol implementation conformance statement (PICS) proforma.

A detailed description of the symbols used in the PICS proforma, along with instructions for completing the PICS proforma, can be found in Clause 21.

92.13.2 Identification

92.13.2.1 Implementation identification

Supplier[1]	
Contact point for enquiries about the PICS[1]	
Implementation Name(s) and Version(s)[1,3]	
Other information necessary for full identification—e.g., name(s) and version(s) for machines and/or operating systems; System Name(s)[2]	
NOTE 1— Required for all implementations. NOTE 2— May be completed as appropriate in meeting the requirements for the identification. NOTE 3—The terms Name and Version should be interpreted appropriately to correspond with a supplier's terminology (e.g., Type, Series, Model).	

92.13.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3bj-20XX, Clause 92, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-CR4
Identification of amendments and corrigenda to this PICS proforma that have been completed as part of this PICS	
Have any Exception items been required? No [ ] Yes [ ] (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3bj-20XX.)	

Date of Statement	

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

Major capabilities/options

Editor's note (to be removed prior to final publication):

Major capabilities/options and PICS proforma tables will be generated when the content of the clause stabilizes.

93. Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4

93.1 Overview

This clause specifies the 100GBASE-KR4 PMD and baseband medium. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 93–1, to the medium through the MDI and to the management functions that are optionally accessible through the management interface defined in Clause 45, or equivalent.

When forming a complete 100GBASE-KR4 Physical Layer, the following guidelines apply.

a) The PMA specified in Clause 83 shall connect to the RS-FEC specified in Clause 91 using the PMA service interface specified in 83.3 with *p*=4 input lanes.
b) The PMA specified in Clause 83 shall connect to the PMD using the service interface specified in 93.2.
c) If one or more CAUI interfaces are implemented between the RS and the RS-FEC, each CAUI interface shall include a pair of PMA sub-layers, as specified in Clause 83. Refer to 83.1.4 for additional guidance.

Table 93–1—Physical Layer clauses associated with the 100GBASE-KR4 PMD

Associated clause	100GBASE-KR4
81—RS	Required
81—CGMII[a]	Optional
82—PCS for 100GBASE-R	Required
91—RS-FEC	Required
83—PMA for 100GBASE-R	Required
83A—CAUI	Optional
73—Auto-Negotiation	Required
78—Energy-Efficient Ethernet	Optional

[a]The CGMII is an optional interface. However, if the CGMII is not implemented, a conforming implementation must behave functionally as though the RS and CGMII were present.

A 100GBASE-KR4 PHY with the optional Energy-Efficient Ethernet (EEE) capability may optionally enter the Low Power Idle (LPI) mode to conserve energy during periods of low link utilization.

Figure 93–1 shows the relationship of the 100GBASE-KR4 PMD sublayers and MDI to the ISO/IEC Open System Interconnection (OSI) reference model.

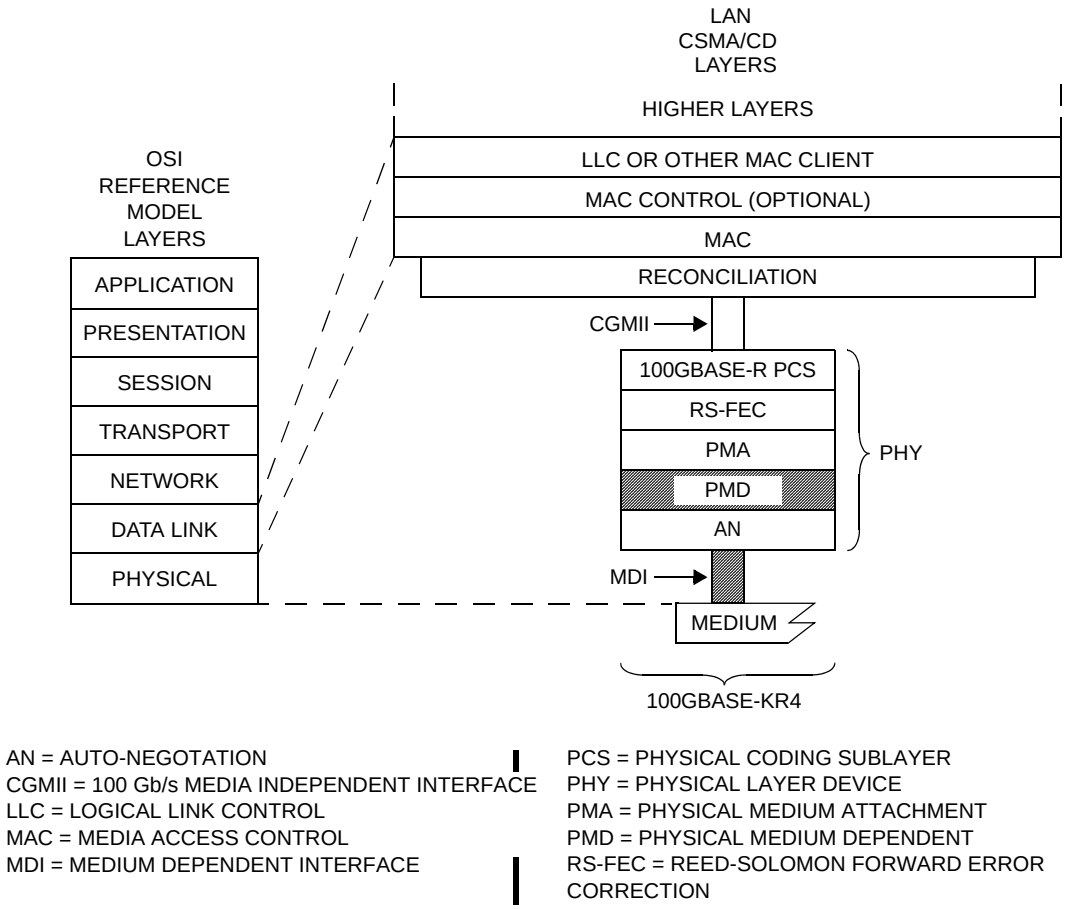


Figure 93–1—100GBASE-KR4 PMD relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

93.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 100GBASE-KR4 PMD. The service interface for this PMD is described in an abstract manner and does not imply any particular implementation. The PMD service interface supports the exchange of encoded data. The PMD translates the encoded data to and from signals suitable for the medium.

The PMD service interface is an instance of the inter-sublayer service interface defined in 80.3. The PMD service interface primitives are summarized as follows:

PMD:IS_UNITDATA_*i*.request
PMD:IS_UNITDATA_*i*.indication
PMD:IS_SIGNAL.indication

The 100GBASE-KR4 PMD has four parallel bit streams, hence $i = 0$ to 3. The PMA (or the PMD) continuously sends four parallel bit streams to the PMD (or the PMA), one per lane, each at a nominal signaling rate of 25.78125 GBd.

SIGNAL_DETECT in 100GBASE-KR4 indicates the successful completion of the start-up protocol on all four lanes. The SIGNAL_DETECT parameter can take on one of two values: OK or FAIL. When SIGNAL_DETECT = FAIL, the PMD:IS_UNITDATA_*i*.indication parameters are undefined. The SIGNAL_DETECT parameter maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

If the optional EEE capability is supported, then the PMD service interface includes two additional primitives as follows:

PMD:IS_TX_MODE.request
PMD:IS_RX_MODE.indicate

Editor's note (to be removed prior to final publication):

The PMD service interface has been augmented as implied by barrass_01_0312 per the motion passed in the March Task Force meeting.

"Move that the 802.3bj Task Force adopt barrass_01_0312.pdf as a baseline for optional Energy Efficient Ethernet operation for 100G Backplane and Twinaxial cable PHYs. (M: M. Gustlin, S: M. Brown, Y: 54, N: 0, A: 3)"

Per barrass_01_0312, PMD transmit function behavior changes as a function of the tx_mode parameter.

DATA, SLEEP, or WAKE: normal behavior
ALERT: send the ALERT signal
FW: send a PMA-specific pattern that is TBD (PMD normal behavior?)
QUIET: disable the transmitter

Per barrass_01_0312, the PMD receive function infers rx_mode from the received signal.

Normal or rapid alignment markers: DATA, SLEEP, or WAKE (mapping not defined)
No signal: QUIET
Alert signal: ALERT
PMA-specific pattern for FW that is TBD: FW (PMA function?)

Given the ambiguity surrounding PMD behavior and the TBD signaling for ALERT, FW, etc., these functions have not been added to the clause. The functions will be added to the clause upon adoption of a detailed proposal describing the desired behavior.

93.3 PCS requirements for Auto-Negotiation (AN) service interface

The PCS associated with this PMD is required to support the AN service interface primitive AN_LINK.indication defined in 73.9. (See 82.6.)

93.4 Delay constraints

The sum of the transmit and the receive delays at one end of the link contributed by the 100GBASE-KR4 PMD, AN, and the medium in one direction shall be no more than TBD bit times (TBD pause_quanta or TBD ns). It is assumed that the one way delay through the medium is no more than TBD bit times (TBD ns).

A description of overall system delay constraints and the definitions for bit times and pause_quanta can be found in 80.4 and its references.

93.5 Skew constraints

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. 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 80.5 and specified at the points SP1 to SP6 shown in Figure 80–4 and Figure 80–5.

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 TBD ns and the Skew Variation at SP2 is limited to TBD ps.

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

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

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 TBD ns and the Skew Variation at SP5 shall be less than TBD ns.

For more information on Skew and Skew Variation see 80.5.

93.6 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that provide control and status information for and about the PMD. If MDIO is implemented, it shall map MDIO control variables to PMD control variables as shown in Table 93–2, and MDIO status variables to PMD status variables as shown in Table 93–3.

Table 93–2—100GBASE-KR4 MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable	1.9.0	Global_PMD_transmit_disable
PMD transmit disable 3 to PMD transmit disable 0	PMD transmit disable	1.9.4 to 1.9.1	PMD_transmit_disable_3 to PMD_transmit_disable_0
Restart training	BASE-R PMD control	1.150.0	mr_restart_training
Training enable	BASE-R PMD control	1.150.1	mr_training_enable

Table 93–3—100GBASE-KR4 MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMD status variable
Fault	PMA/PMD status 1	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect	1.10.0	Global_PMD_signal_detect
PMD receive signal detect 3 to PMD receive signal detect 0	PMD receive signal detect	1.10.4 to 1.10.1	PMD_signal_detect_9 to PMD_signal_detect_0
Receiver status 3	BASE-R PMD status	1.151.12	rx_trained_3
Frame lock 3	BASE-R PMD status	1.151.13	frame_lock_3
Start-up protocol status 3	BASE-R PMD status	1.151.14	training_3
Training failure 3	BASE-R PMD status	1.151.15	training_failure_3
Receiver status 2	BASE-R PMD status	1.151.8	rx_trained_2
Frame lock 2	BASE-R PMD status	1.151.9	frame_lock_2
Start-up protocol status 2	BASE-R PMD status	1.151.10	training_2
Training failure 2	BASE-R PMD status	1.151.11	training_failure_2
Receiver status 1	BASE-R PMD status	1.151.4	rx_trained_1
Frame lock 1	BASE-R PMD status	1.151.5	frame_lock_1
Start-up protocol status 1	BASE-R PMD status	1.151.6	training_1
Training failure 1	BASE-R PMD status	1.151.7	training_failure_1
Receiver status 0	BASE-R PMD status	1.151.0	rx_trained_0
Frame lock 0	BASE-R PMD status	1.151.1	frame_lock_0
Start-up protocol status 0	BASE-R PMD status	1.151.2	training_0
Training failure 0	BASE-R PMD status	1.151.3	training_failure_0

93.7 PMD functional specifications

93.7.1 Link block diagram

One direction from one lane of a 100GBASE-KR4 link is shown in Figure 93–2.

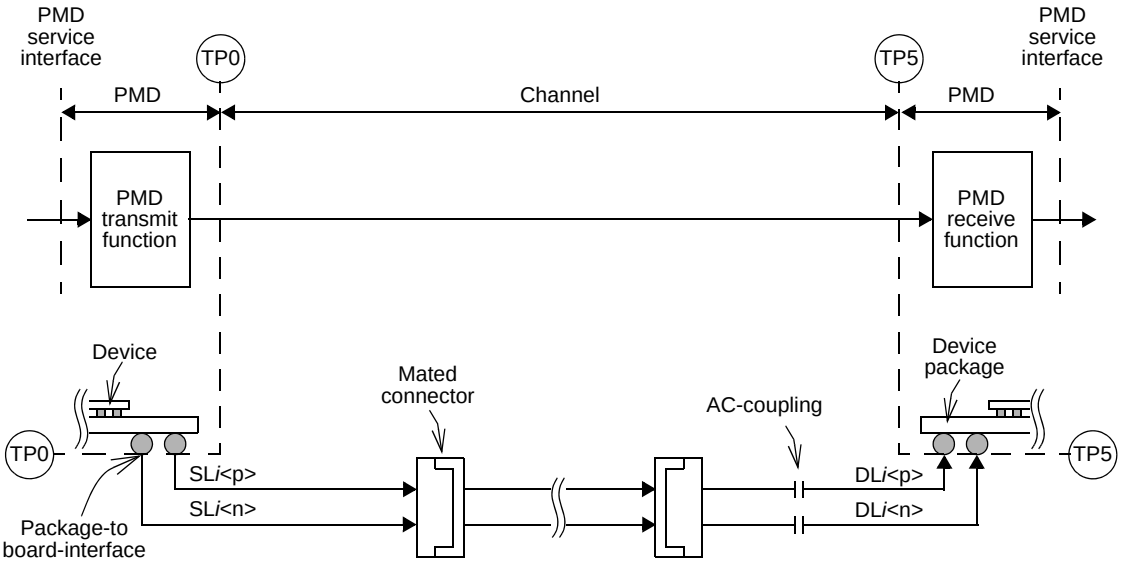


Figure 93–2—100GBASE-KR4 link (one direction from one lane is illustrated)

93.7.2 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 electrical streams. The four electrical signal streams shall then be delivered to the MDI, all according to the transmit electrical specifications in 93.8.1. A positive output voltage of SL*i*<p> minus SL*i*<n> (differential voltage) shall correspond to tx_bit = one.

93.7.3 PMD Receive function

The PMD receive function shall convert the four electrical streams from the MDI 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. A positive input voltage of DL*i*<p> minus DL*i*<n> (differential voltage) shall correspond to rx_bit = one.

93.7.4 Global PMD signal detect function

The Global PMD signal detect function shall continuously report the message PMD:IS_SIGNAL.indication (SIGNAL_DETECT) to the PMD service interface. SIGNAL_DETECT, while normally intended to be an indicator of signal presence, is used by the 100GBASE-KR4 PMD to indicate the successful completion of the start-up protocol on all lanes. The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

SIGNAL_DETECT shall be set to FAIL following system reset or the manual reset of the training state diagram. Upon successful completion of training on all lanes, SIGNAL_DETECT shall be set to OK.

If training is disabled by management, SIGNAL_DETECT shall be set to OK.

If the MDIO interface is implemented, then Global_PMD_signal_detect (1.10.0) shall be continuously set to the value of SIGNAL_DETECT as described in 45.2.1.9.7.

93.7.5 PMD lane-by-lane signal detect function

When the MDIO is implemented, each PMD_signal_detect_*i* value, where *i* represents the lane number in the range 0 to 3, shall be continuously updated as described in the following two paragraphs.

PMD_signal_detect_0 (1.10.1), PMD_signal_detect_1 (1.10.2), PMD_signal_detect_2 (1.10.3), and PMD_signal_detect_3 (1.10.4) shall be set to one or zero depending on whether a particular lane's signal_detect, as defined by the training state diagram in Figure 72–5, returns true or false.

93.7.6 Global PMD transmit disable function

The Global_PMD_transmit_disable function is optional. When implemented, it allows all of the transmitters to be disabled with a single variable.

a) When Global_PMD_transmit_disable variable is set to one, this function shall turn off all of the transmitters such that each transmitter drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage in Table 93–4.
b) If a PMD_fault (93.7.9) is detected, then the PMD may turn off the electrical transmitter in all lanes.
c) Loopback, as defined in 93.7.8, shall not be affected by Global_PMD_transmit_disable.

93.7.7 PMD lane-by-lane transmit disable function

The PMD_transmit_disable_*i* function (where *i* represents the lane number in the range 0:3) is optional and allows the electrical transmitter in each lane to be selectively disabled. When this function is supported, it shall meet the following requirements:

a) When a PMD_transmit_disable_*i* variable is set to one, this function shall turn off the transmitter associated with that variable such that it drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage specified in Table 93–4.
b) If a PMD_fault (93.7.9) is detected, then the PMD may set each PMD_transmit_disable_*i* to one, turning off the electrical transmitter in each lane.
c) Loopback, as defined in 93.7.8, shall not be affected by PMD_transmit_disable_*i*.

93.7.8 Loopback mode

Local loopback mode shall be provided by the adjacent PMA (see 83.5.8) as a test function to the device. When loopback mode is enabled, transmission requests passed to each transmitter are sent directly to the corresponding receiver, overriding any signal detected by each receiver on its attached link. Note that loopback mode does not affect the state of the transmitter, which continues to send data (unless disabled).

Control of the loopback function is specified in 45.2.1.1.5.

NOTE 1—The signal path that is exercised in the loopback mode is implementation specific, but it is recommended that this signal path encompass as much of the circuitry as is practical. The intention of providing this loopback mode of operation is to permit diagnostic or self-test functions to test the transmit and receive data paths using actual data. Other loopback signal paths may also be enabled independently using loopback controls within other devices or sublayers.

NOTE 2—Placing a network port into loopback mode can be disruptive to a network.

93.7.9 PMD_fault function

If the MDIO is implemented, PMD_fault is the logical OR of PMD_receive_fault, PMD_transmit_fault, and any other implementation specific fault.

93.7.10 PMD transmit fault function

The PMD_transmit_fault function is optional. The faults detected by this function are implementation specific, but should not include the assertion of the Global_PMD_transmit_disable function.

If a PMD_transmit_fault (optional) is detected, then the Global_PMD_transmit_disable function should also be asserted.

If the MDIO interface is implemented, then this function shall be mapped to the Transmit fault bit as specified in 45.2.1.7.4.

93.7.11 PMD receive fault function

The PMD_receive_fault function is optional. The faults detected by this function are implementation specific.

If the MDIO interface is implemented, then this function shall contribute to the Receive fault bit as specified in 45.2.1.7.5.

93.7.12 PMD control function

Each lane of the 100GBASE-KR4 PMD shall use the same control function as 10GBASE-KR, as defined in 72.6.10.

The random seed for the training pattern described in 72.6.10.2.6 shall be different for each of the lanes.

The variables rx_trained_*i*, frame_lock_*i*, training_*i*, and training_failure_*i* (where *i* goes from 0 to 3) report status for each lane and are equivalent to rx_trained, frame_lock, training, and training_failure as defined in 72.6.10.3.1.

If the MDIO interface is implemented, then this function shall map these variables to the appropriate bits in the BASE-R PMD status register (Register 1.151) as specified in 45.2.1.80.

93.8 100GBASE-KR4 electrical characteristics

93.8.1 Transmitter characteristics

Transmitter characteristics measured at TP0a are summarized in Table 93–4.

93.8.1.1 Transmitter test fixture

Unless otherwise noted, measurements of the transmitter are made at the output of a test fixture as shown in Figure 93–3.

The insertion loss of the test fixture shall be between 1.3 dB and 1.6 dB at 12.89 GHz. The return loss of the text fixture shall be less than or equal to 12 dB from 0.05 GHz to 13 GHz.

Table 93–4—Summary of transmitter characteristics at TP0a

Parameter	Subclause reference	Value	Units
Signaling rate	93.8.1.2	25.78125±100 ppm	GBd
Differential peak-to-peak output voltage (max.) Transmitter disabled[a] Transmitter enabled	93.8.1.3	 30 1200	 mV mV
DC common-mode output voltage (max.)	93.8.1.3	1.9	V
DC common-mode output voltage (min.)	93.8.1.3	0	V
AC common-mode output voltage (RMS, max.)	93.8.1.3	12	mV
Differential output return loss (min.)	93.8.1.4	Equation (93–1)	dB
Common-mode output return loss (min.)	93.8.1.4	Equation (93–2)	dB
Transition time (20-80%, min.), no equalization[b]	93.8.1.5	8	ps
Output waveform Steady-state voltage v_f (max.) Steady-state voltage v_f (min.) Linear fit pulse peak (min.) Normalized RMS linear fit error (max.) Normalized coefficient step size (min.) Normalized coefficient step size (max.) Pre-cursor full-scale range (min.) Post-cursor full-scale range (min.)	93.8.1.6	 0.6 0.4 0.8 × v_f 0.037 0.0083 0.05 1.54 4	 V V V — — — — —
Far-end output noise (max.) Low insertion loss channel High insertion loss channel	93.8.1.7	 2 1	 mV mV
Output jitter (max.) Effective random jitter Even-odd jitter Total jitter excluding data dependent jitter	93.8.1.8	 0.15 0.035 0.28	 UI UI UI

[a]The transmitter for lane *i* is disabled when either Global_PMD_transmit_disable or PMD_transmit_disable_*i* is set to one.
[b]Transmit equalization may be disabled by asserting the preset control defined in Table 45–60 and 45.2.1.81.3.

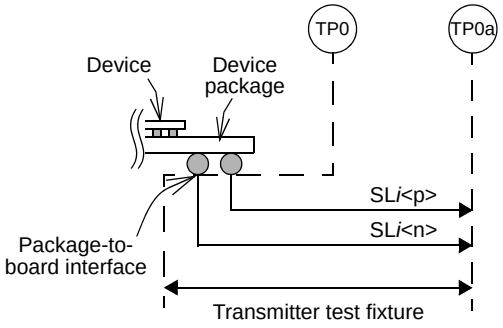


Figure 93–3—Transmitter test fixture and test points

93.8.1.2 Signaling rate and range

The 100GBASE-KR4 signaling rate shall be 25.78125 GBd ± 100 ppm per lane.

93.8.1.3 Signal levels

The differential output voltage v_{di} is defined to be SL*i*<p> minus SL*i*<n>. The common-mode output 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 93–4.

For a square wave test pattern with a 2 UI period, the peak-to-peak differential output voltage shall be less than or equal to 1200 mV regardless of the transmit equalizer setting. The peak-to-peak differential output voltage shall be less than or equal to 30 mV when the transmitter is disabled (refer to 93.7.6 and 93.7.7). The differential output voltage test pattern shall consist of no fewer than eight symbols of alternating polarity.

The DC common-mode output voltage shall be between 0 V and 1.9 V with respect to signal ground. The AC common-mode output voltage shall be less than or equal to 12 mV RMS with respect to signal ground.

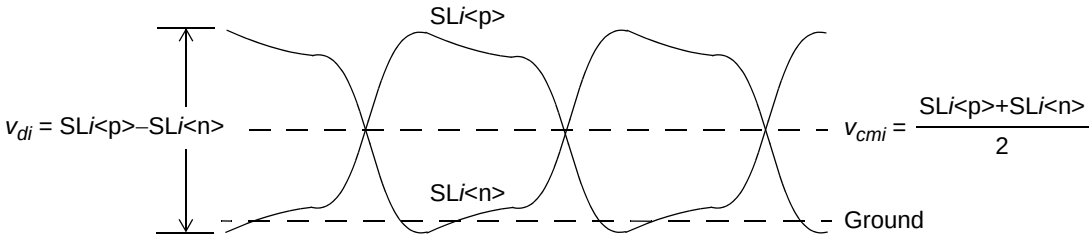


Figure 93–4—Transmitter output voltage definitions

93.8.1.4 Transmitter output return loss

The differential output return loss, in dB, of the transmitter shall meet Equation (93–1). This output impedance requirement applies to all valid output levels. The reference impedance for differential return loss measurements shall be 100 Ω.

TBD (93–1)

The common-mode output return loss, in dB, of the transmitter shall meet Equation (93–2). This output impedance requirement applies to all valid output levels. The reference impedance for common-mode return loss measurements shall be 25 Ω.

TBD (93–2)

93.8.1.5 Transition time

Editor's note (to be removed prior to final publication):

The following definition for transition time is based on 86A.5.3.3.

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 (see 68.6.2).

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 zeros. In this case, the 0% level and the 100% level may be estimated as ZeroLevel and ZeroLevel+MeasuredOMA in the TWDP code (see 68.6.6.2) or as the average signal within windows from –3 UI to –2 UI and from 2 UI to 3 UI relative to the edge.

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.

The transition times shall be greater than or equal to 8 ps when transmit equalization is disabled. Transmit equalization may be disabled by asserting the preset control defined in Table 45–60 and 45.2.1.81.3.

93.8.1.6 Transmitter output waveform

The 100GBASE-KR4 transmit function includes programmable equalization to compensate for the frequency-dependent loss of the channel and facilitate data recovery at the receiver. The functional model for the transmit equalizer is the three tap transversal filter shown in Figure 93–5.

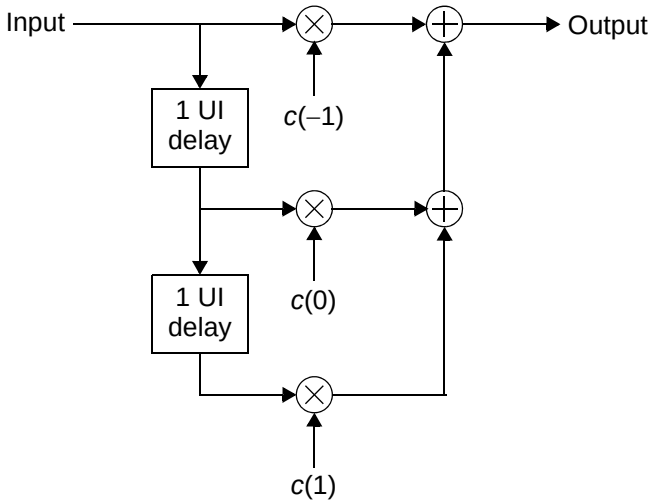


Figure 93–5—Transmit equalizer functional model

The state of the transmit equalizer and hence the transmitted output waveform may be manipulated via the PMD control function defined in 93.7.12 or via the management interface. The transmit function responds to a set of commands issued by the link partner's receive function and conveyed by a back-channel communications path.

This command set includes instructions to:

Increment coefficient $c(i)$.
Decrement coefficient $c(i)$.
Hold coefficient $c(i)$ at its current value.
Set the coefficients to a pre-defined value (preset or initialize).

In response, the transmit function relays status information to the link partner's receive function. The status messages indicate that:

The requested update to coefficient $c(i)$ has completed (updated).
Coefficient $c(i)$ is at its minimum value.
Coefficient $c(i)$ is at its maximum value.
Coefficient $c(i)$ is ready for the next update request (not_updated).

The transmitter output waveform is characterized using the procedure described in 85.8.3.3. The parameters of the linear pulse fit and equalizer are summarized in Table 93–5.

Table 93–5—Linear fit pulse and equalizer parameters

Description	Symbol	Value	Units
Linear fit pulse length	N_p	8	UI
Linear fit pulse delay	D_p	2	UI
Equalizer length	N_w	8	UI
Equalizer delay	D_w	2	UI

93.8.1.6.1 Steady state voltage and linear fit pulse peak

The steady state voltage and linear fit pulse peak values shall be verified after the transmit equalizer coefficients have been set to the “preset” values.

The steady state voltage v_f is defined to be the sum of the linear fit pulse $p(k)$ divided by M (refer to 85.8.3.3 step 3). The steady state voltage shall be greater than or equal to 0.4 V and less than or equal to 0.6 V.

The peak value of $p(k)$ shall be greater than $0.8 \times v_f$.

93.8.1.6.2 Linear fit error

For any configuration of the transmit equalizer, the RMS value of the error between the linear fit and the measured waveform, $e(k)$, normalized to the peak value of the linear fit pulse, $p(k)$, shall be less than or equal to 0.037.

93.8.1.6.3 Coefficient initialization

When the PMD enters the INITIALIZE state of the Training state diagram (Figure 72–5) or receives a valid request to “initialize” from the link partner, the coefficients of the transmit equalizer shall be configured such that the ratio $(c(0)+c(1)–c(–1))/(c(0)+c(1)+c(–1))$ is TBD and the ratio $(c(0)–c(1)+c(–1))/(c(0)+c(1)+c(–1))$ is TBD. These requirements apply upon the assertion a coefficient status report of “updated” for all coefficients.

93.8.1.6.4 Coefficient step size

The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to “increment” that coefficient shall be between 0.0083 and 0.05. The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to “decrement” that coefficient shall be between –0.05 and –0.0083.

The change in the normalized amplitude of the coefficient is defined to be the difference in the value measured prior to the assertion of the “increment” or “decrement” request (e.g., the coefficient update request for

all coefficients is “hold”) and the value upon the assertion of a coefficient status report of “updated” for that coefficient.

93.8.1.6.5 Coefficient range

When sufficient “increment” or “decrement” requests have been received for a given coefficient, the coefficient will reach a lower or upper bound based on the coefficient range or restrictions placed on the minimum steady state differential output voltage or the maximum peak-to-peak differential output voltage.

With $c(-1)$ set to zero and both $c(0)$ and $c(1)$ having received sufficient “decrement” requests so that they are at their respective minimum values, the ratio $(c(0)–c(1))/(c(0)+c(1))$ shall be greater than or equal to 4.

With $c(1)$ set to zero and both $c(-1)$ and $c(0)$ having received sufficient “decrement” requests so that they are at their respective minimum values, the ratio $(c(0)–c(-1))/(c(0)+c(-1))$ shall be greater than or equal to 1.54.

Note that a coefficient may be set to zero by first asserting the preset control and then manipulating the other coefficients as required by the test.

93.8.1.7 Transmitter far-end output noise

The transmitter far-end output noise is a source of noise in addition to the channel integrated crosstalk noise (ICN) specified in 93.9.3. The transmitter far-end output noise is characterized using the procedure defined in 85.8.3.2.

The far-end transmitter output noise is characterized using two reference channels:

A “low-loss” channel with TBD dB insertion loss at 12.89 GHz and ICN that meets the requirements of 93.9.3.
A “high-loss” channel with TBD dB insertion loss at 12.89 GHz and ICN that meets the requirements of 93.9.3.

For the low-loss channel, the far-end transmitter output noise shall be less than or equal to 2 mV RMS. For the high-loss channel, the far-end transmitter output noise shall be less than or equal to 1 mV RMS.

93.8.1.8 Transmitter output jitter

Even-odd jitter is characterized using the procedure defined in 92.8.3.8. Even-odd jitter shall be less than or equal to 0.035 UI regardless of the transmit equalization setting.

Total jitter is characterized using the procedure defined in 92.8.3.8. Data dependent jitter is characterized using the procedure defined in 85.8.3.8. The total jitter, excluding data dependent jitter, shall be less than or equal to 0.28 UI regardless of the transmit equalization setting.

Editor’s note (to be removed prior to final publication):

Clause 85 was not clear on what was meant by “excluding” DDJ from TJ. In this draft, it has been interpreted to be the value of TJ as defined in 92.8.3.8 minus the value of DDJ as defined in 85.8.3.8.

The effective random jitter is characterized using the procedure defined in 92.8.3.8. The effective random jitter shall be less than or equal to 0.15 UI regardless of the transmit equalization setting.

Editor's note (to be removed prior to final publication):

Clause 72 specifies that transmit equalization is off (preset) for the measurement of jitter. Clause 85 does not specify the conditions for jitter measurement and this is interpreted to mean that it applies for all transmit equalization settings. There has been some discussion as to which approach is appropriate.

93.8.2 Receiver characteristics

Receiver characteristics measured at TP5a are summarized in Table 93–6.

Table 93–6—Summary of receiver characteristics at TP5a

Parameter	Subclause reference	Value	Units
Differential input return loss (min.)	93.8.2.2	Equation (93–3)	dB
Common-mode input return loss (min.)	93.8.2.2	Equation (93–4)	dB
Differential to common-mode return loss (min.)	93.8.2.2	Equation (93–5)	dB

93.8.2.1 Receiver test fixture

Unless otherwise noted, measurements of the receiver are made at the input to a test fixture as shown in Figure 93–6.

The insertion loss of the test fixture shall be between 1.3 dB and 1.6 dB at 12.89 GHz. The return loss of the text fixture shall be less than or equal to 12 dB from 0.05 GHz to 13 GHz.

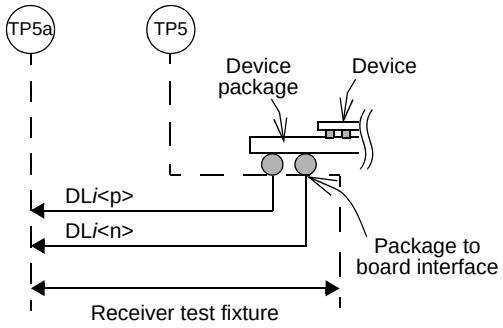


Figure 93–6—Receiver test fixture and test points

93.8.2.2 Receiver input return loss

The differential input return loss, in dB, of the receiver shall meet Equation (93–3). The reference impedance for differential return loss measurements shall be 100 Ω.

TBD (93–3)

The common-mode input return loss, in dB, of the receiver shall meet Equation (93–4). The reference impedance for common-mode return loss measurements shall be 25 Ω.

TBD (93–4)

The differential to common-mode return loss, in dB, of the receiver shall meet Equation (93–5).

TBD (93–5)

93.8.2.3 Receiver interference tolerance

Editor's note (to be removed prior to final publication):

It is not clear whether or not 85.8.4.2 is the correct reference for the interference tolerance test procedure but it will be used pending a more detailed proposal. Another likely reference point in Annex 69A.

Receiver interference tolerance is characterized using the procedure defined in 85.8.4.2. The receiver shall satisfy the requirements for interference tolerance summarized in Table 93–7.

Table 93–7—Receiver interference tolerance parameters

Parameter	Test 1 values	Test 2 values	Test 3 values	Test 4 values	Units
Maximum BER without FEC	10^{-12}	10^{-12}	2×10^{-5}	2×10^{-5}	
Channel insertion loss at 12.89 GHz	TBD	TBD	TBD	TBD	dB
Real part of a_0, min.[a]	TBD	TBD	TBD	TBD	
Real part of a_1, min.	TBD	TBD	TBD	TBD	$Hz^{-1/2}$
Real part of a_2, min.	TBD	TBD	TBD	TBD	Hz^{-1}
Real part of a_4, min.	TBD	TBD	TBD	TBD	Hz^{-2}
Channel noise excluding TX-RX re-reflection noise	0	0	0	0	mV
Applied peak-to-peak Sinusoidal Jitter[b]	TBD	TBD	TBD	TBD	UI
Applied peak-to-peak Random Jitter[c]	TBD	TBD	TBD	TBD	UI
Applied Duty Cycle Distortion	TBD	TBD	TBD	TBD	UI
Applied RMS broadband noise	TBD	TBD	TBD	TBD	mV

[a]For each test channel, a_0 is limited to a maximum value of 1.5 and a_1, a_2, and a_4 are limited to a maximum value of 0.
[b]The frequency of the sinusoid must be greater than 100 MHz.
[c]Random Jitter is specified at a BER of 10^{-12}.

93.8.3 AC-coupling

The 100GBASE-KR4 transmitter shall be AC-coupled to the receiver. Common-mode specifications are defined as if the DC-blocking capacitor is implemented between TP0 and TP5. Should the capacitor be implemented outside TP0 and TP5, the common-mode specifications in Table 93–4 may not be appropriate.

The impact of a DC-blocking capacitor implemented between TP0 and TP5 is accounted for within the channel specifications. Should the capacitor be implemented outside TP0 and TP5, it is the responsibility of

implementers to consider any necessary modifications to common-mode and channel specifications required for interoperability as well as any impact on the verification of transmitter and receiver compliance.

The low frequency 3 dB cutoff of the AC coupling shall be less than 50 kHz.

93.9 Channel characteristics

93.9.1 Channel operating margin

The channel operating margin (COM) computed using the procedure in Annex 93A and the parameters in Table 93–8 shall be greater than or equal to TBD dB.

93.9.2 Insertion loss

The insertion loss, in dB, of the channel is recommend to meet Equation (93–6).

$$IL(f) \le \begin{cases} 1.50 + 4.60\sqrt{f} + 1.318f & 0.05 \le f \le f_b/2 \\ -12.71 + 3.70f & f_b/2 < f \le f_b \end{cases} \quad (93\text{–}6)$$

where

f is the frequency in GHz
f_b is the signaling rate (25.78125) in GHz
$IL(f)$ is the insertion loss at frequency f

The insertion loss limit is illustrated by Figure 93–7.

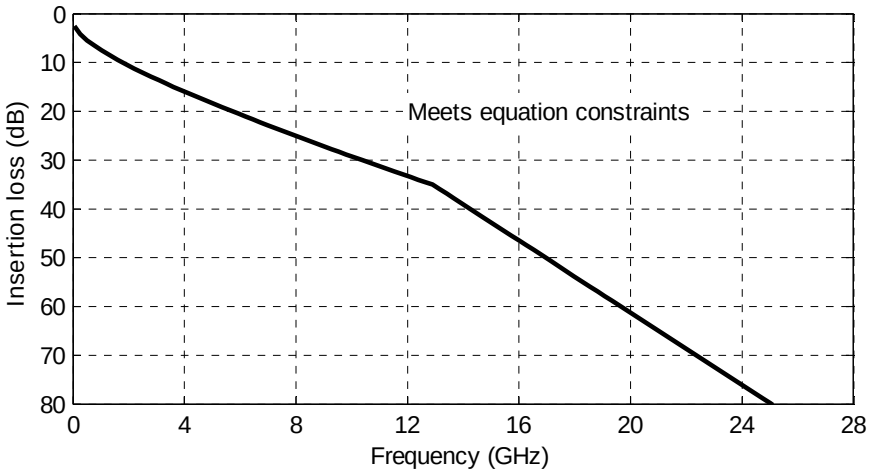


Figure 93–7—Insertion loss limit

Table 93–8—Channel operating margin parameters

Parameter	Symbol	Value	Units
Signaling rate	f_b	25.78125	GHz
Maximum start frequency	f_{min}	0.05	GHz
Maximum frequency step	Δf	0.01	GHz
Transmitter reflection coefficient DC value Reference frequency	Γ_1 Γ_{01} f_1	 TBD TBD	 V/V GHz
Receiver reflection coefficient DC value Reference frequency	Γ_2 Γ_{02} f_2	 TBD TBD	 V/V GHz
Transmitter differential peak output voltage Victim Far-end aggressor Near-end aggressor	 A_v A_f A_n	 0.4 0.4 0.6	 V V V
Transmitter 3 dB bandwidth Victim Far-end aggressor Near-end aggressor	 f_v f_f f_n	 $0.55 \times f_b$ $0.55 \times f_b$ f_b	 GHz GHz GHz
Receiver 3 dB bandwidth	f_r	$0.75 \times f_b$	
Transmitter equalizer, pre-cursor coefficient Minimum value Maximum value Step size	$c(-1)$	 –0.10 0.00 0.02	 — — —
Transmitter equalizer, post-cursor coefficient Minimum value Maximum value Step size	$c(1)$	 –0.40 0.00 0.02	 — — —
Continuous time filter, DC gain Minimum value Maximum value Step size	g_{DC}	 –12 0 1	 dB dB dB
Number of signal levels	L	2	—
Number of samples per unit interval	M	32	—
Victim single bit response exception window	W	TBD	UI
Normalized RMS Gaussian noise	σ_G	0.01	—
Normalized peak dual-Dirac noise	A_{DD}	0.1	—
Target uncorrected symbol error ratio	SER_0	10^{-5}	—
Minimum channel operating margin	COM_0	TBD	dB

93.9.3 Channel Integrated Crosstalk Noise (ICN)

Editor's note (to be removed prior to final publication):

This is a placeholder to be referenced by the transmitter far-end output noise specification (refer to 93.8.1.7).

93.10 Environmental specifications

93.10.1 General safety

All equipment subject to this clause shall conform to applicable sections (including isolation requirements) of IEC 60950-1.

93.10.2 Network safety

The designer is urged to consult the relevant local, national, and international safety regulations to ensure compliance with the appropriate requirements.

93.10.3 Installation and maintenance guidelines

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

93.10.4 Electromagnetic compatibility

A system integrating the 100GBASE-KR4 PHY shall comply with applicable local and national codes for the limitation of electromagnetic interference.

93.10.5 Temperature and humidity

A system integrating the 100GBASE-KR4 PHY is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration). Specific requirements and values for these parameters are considered to be beyond the scope of this standard.

93.11 Protocol implementation conformance statement (PICS) proforma for Clause 93, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4[11]

93.11.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 93, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4, shall complete the following protocol implementation conformance statement (PICS) proforma.

A detailed description of the symbols used in the PICS proforma, along with instructions for completing the PICS proforma, can be found in Clause 21.

93.11.2 Identification

93.11.2.1 Implementation identification

Supplier[1]	
Contact point for enquiries about the PICS[1]	
Implementation Name(s) and Version(s)[1,3]	
Other information necessary for full identification—e.g., name(s) and version(s) for machines and/or operating systems; System Name(s)[2]	
NOTE 1— Required for all implementations. NOTE 2— May be completed as appropriate in meeting the requirements for the identification. NOTE 3—The terms Name and Version should be interpreted appropriately to correspond with a supplier's terminology (e.g., Type, Series, Model).	

93.11.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3bj-20XX, Clause 93, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KR4
Identification of amendments and corrigenda to this PICS proforma that have been completed as part of this PICS	
Have any Exception items been required? No [ ] Yes [ ] (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3bj-20XX.)	

Date of Statement	

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

93.11.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
CGMII	CGMII	93.1	Interface is supported	O	Yes [ ] No [ ]
PCS	100GBASE-R PCS	93.1		M	Yes [ ]
FEC	Forward error correction	93.1		TBD	Yes [ ] No [ ]
PMA	100GBASE-R PMA	93.1		M	Yes [ ]
CAUI	CAUI	93.1	Interface is supported	O	Yes [ ] No [ ]
AN	Auto-negotiation	93.1		M	Yes [ ]
DC	Delay constraints	93.4		M	Yes [ ]
DSC	Skew constraints	93.5		M	Yes [ ]
*MD	MDIO capability	93.6	Registers and interface supported	O	Yes [ ] No [ ]

93.11.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and base-band medium, type 100GBASE-KR4

Editor's note (to be removed prior to final publication):

PICS proforma tables will be generated when the content of the clause stabilizes.

94. Physical Medium Attachment (PMA) sublayer, Physical Medium Dependent (PMD) sublayer, and baseband medium, type 100GBASE-KP4

94.1 Overview

This clause specifies the Physical Medium Attachment (PMA) sublayers, Physical Medium Dependent (PMD) sublayer, and medium for the 100GBASE-KP4 PHY. A 100GBASE-KP4 physical shall include the required sub-layers and may include the optional sub-layers specified in Table 94–1.

When forming a complete 100GBASE-KP4 Physical Layer, the following guidelines apply:

a) The 100GBASE-KP4 PMA specified in 94.2 shall connect with the FEC through the PMA service interface specified in 94.2.1.

b) The 100GBASE-KP4 PMD specified in 94.3 shall connect with the 100GBASE-KP4 PMA through the PMD service interface specified in 94.3.1.

c) If one or more CAUI interfaces are implemented between the FEC and the RS, each CAUI interface shall include a pair of PMA sub-layers, as specified in Clause 83. Refer to 83.1.4 for additional guidance.

Table 94–1—Physical Layer clauses associated with the 100GBASE-KP4 PMD

Associated clause	100GBASE-KP4
81—RS	Required
81—CGMII[a]	Optional
82—PCS for 100GBASE-R	Required
83—PMA	Optional
83A—CAUI	Optional
91—RS-FEC	Required
94—PMA	Required
94—PMD	Required
73—Auto-Negotiation	Required
78—Energy-Efficient Ethernet	Optional

[a]The CGMII is an optional interface. However, if the CGMII is not implemented, a conforming implementation must behave functionally as though the RS and CGMII were present.

Figure 94–1 shows the relationship of the 100GBASE-KP4 PMA and PMD sublayers and MDI to the ISO/IEC Open System Interconnection (OSI) reference model.

A 100GBASE-KP4 PHY with the optional Energy-Efficient Ethernet (EEE) capability may optionally enter the Low Power Idle (LPI) mode to conserve energy during periods of low link utilization.

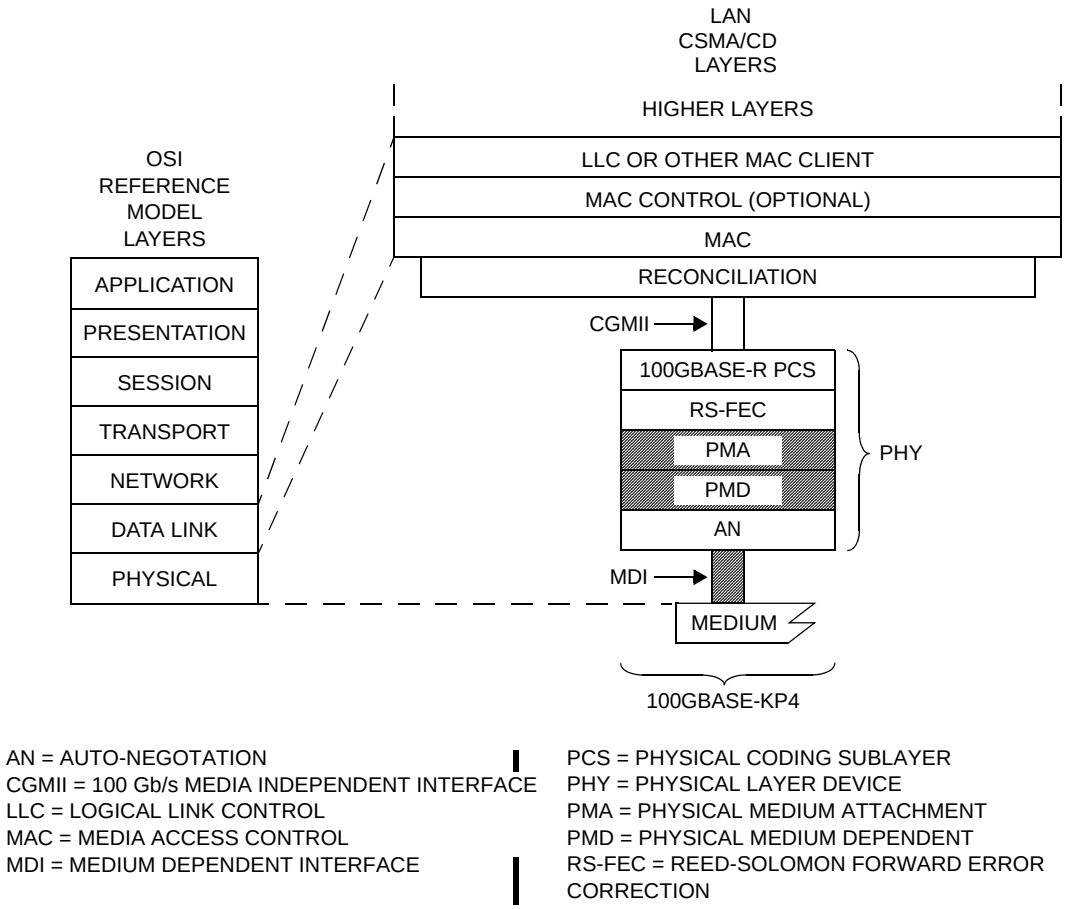


Figure 94–1—100GBASE-KP4 PMA and PMD relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

94.2 Physical Medium Attachment (PMA) Sublayer

94.2.1 PMA Service Interface

The PMA service interface for 100GBASE-KP4 PMA is based on the inter-sublayer service interface defined in 80.3. This interface is defined in an abstract manner and does not imply any particular implementation.

The PMA service interface primitives are summarized as follows:

PMA:IS_UNITDATA_*i*.request
PMA:IS_UNITDATA_*i*.indication
PMA:IS_SIGNAL.indication

If the optional Energy-Efficient Ethernet (EEE) capability is supported (see 78) then the PMA service interface includes two additional primitives as follows:

PMA:IS_TX_MODE.request
PMA:IS_RX_MODE.request

Editor's note (to be removed prior to final publication):

Per barrass_01_0312, PMA transmit function behavior changes as a function of the tx_mode parameter.

DATA, SLEEP, or WAKE: normal behavior
ALERT: send the ALERT signal
FW: send a PMA-specific pattern that is TBD (PMD normal behavior?)
QUIET: disable the transmitter

Per barrass_01_0312, the PMA receive function infers rx_mode from the received signal.

Normal or rapid alignment markers: DATA, SLEEP, or WAKE (mapping not defined)
No signal: QUIET
Alert signal: ALERT
PMA-specific pattern for FW that is TBD: FW (PMA function?)

Given the ambiguity surrounding PMA behavior and the TBD signaling for ALERT, FW, etc., these functions have not been added to the clause. The functions will be added to the clause upon adoption of a detailed proposal describing the desired behavior.

94.2.1.1 PMA:IS_UNITDATA_i.request

The PMA:IS_UNITDATA_*i*.request (where *i*=0 to 3) primitive is used to define the transfer of four streams of data units from the PMA client to PMA.

94.2.1.1.1 Semantics of the service primitive

PMA:IS_UNITDATA_0.request(tx_bit,start)
PMA:IS_UNITDATA_1.request(tx_bit,start)
PMA:IS_UNITDATA_2.request(tx_bit,start)
PMA:IS_UNITDATA_3.request(tx_bit.start)

The data conveyed by PMA:IS_UNITDATA_0.request to IS_UNITDATA_3.request consists of four parallel continuous streams of encoded bits, one stream for each lane. Each of the tx_bit parameters can take one of two values: one or zero. The start parameter is TRUE to indicate that the concurrent tx_bit is the first bit of the first, second, third, or fourth FEC symbol in the FEC codeword and is otherwise FALSE.

94.2.1.1.2 When generated

The PMA client continuously sends four parallel bit streams PMA:IS_UNITDATA_*i*.request(tx_bit,start) to the PMA, each at a nominal signaling rate of 26.5625 Gb/s.

94.2.1.1.3 Effect of receipt

Upon receiving each instance of PMA:IS_UNITDATA_*i*.request, the tx_bit and start parameters are passed to the PMA framing process corresponding to each stream.

94.2.1.2 PMA:IS_UNITDATA_i.indication

The PMA:IS_UNITDATA_*i*.indication (where *i*=0 to 3) primitive is used to define the transfer of four streams of data units from the PMA to the PMA client.

94.2.1.2.1 Semantics of the service primitive

PMA:IS_UNITDATA_0.indication(rx_bit,start)
PMA:IS_UNITDATA_1.indication(rx_bit,start)
PMA:IS_UNITDATA_2.indication(rx_bit,start)
PMA:IS_UNITDATA_3.indication(rx_bit,start)

The data conveyed by PMA:IS_UNITDATA_0.indication to PMA:IS_UNITDATA_3.indication consists of 4 parallel continuous streams of encoded bits, one stream for each lane. Each of the rx_bit parameters can take one of two values: one or zero. The start parameter is TRUE to indicate that the concurrent rx_bit is the first bit of the first, second, third, or fourth FEC symbol in the FEC codeword and is otherwise FALSE.

94.2.1.2.2 When generated

The PMA continuously sends four parallel bit streams PMA:IS_UNITDATA_*i*.indication(rx_bit,start) to the PMA client, each at a nominal signaling rate of 26.5625 Gb/s.

94.2.1.2.3 Effect of receipt

The effect of receipt of this primitive is defined by the PMA client.

94.2.1.3 PMA:IS_SIGNAL.indication

The PMA:IS_SIGNAL.indication primitive is generated by the PMA to the PMA client to indicate the status of the receive process. This primitive is generated by the receive process to propagate the detection of severe error conditions (e.g., loss of synchronization) to the PMA client.

94.2.1.3.1 Semantics of the service primitive

PMA:IS_SIGNAL.indication(SIGNAL_OK)

The SIGNAL_OK parameter can take on one of two values: OK or FAIL. A value of FAIL denotes that invalid data is being presented (rx_bit parameters undefined) by the PMA to the PMA client. A value of OK does not guarantee valid data is being presented by the PMA to the PMA client.

94.2.1.3.2 When generated

The PMA generates the PMA:IS_SIGNAL.indication primitive to the PMA client whenever there is a change in the value of the SIGNAL_OK parameter.

94.2.1.4 Effect of receipt

The effect of receipt of this primitive is defined by the PMA client.

94.2.2 PMA Transmit Functional Specifications

In the transmit direction, the role of the 100GBASE-KP4 PMA is to adapt the signal from the FEC to a PAM4 encoded signal to be passed to the PMD for transfer over the attached medium. The adaptation processes shown in Figure 94–2 include insert overhead, insert termination bits, apply Gray coding, apply 1/(1+D) mod 4 precoding, and apply PAM4 encoding.

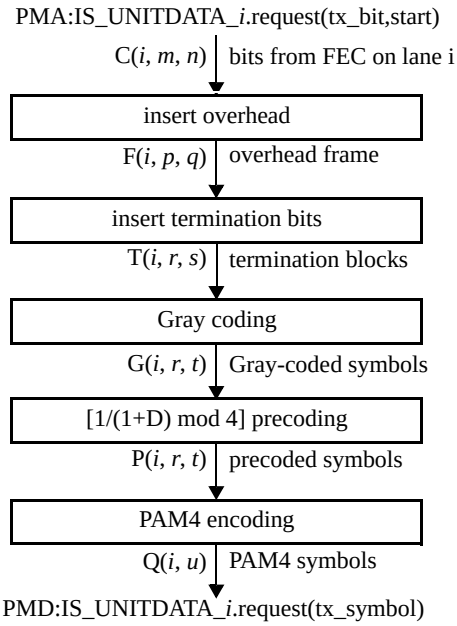


Figure 94–2—Transmit adaptation process diagram

94.2.2.1 FEC Interface

The PMA receives FEC bits via the PMA:IS_UNITDATA_*i*(tx_bit,start) primitive (see 94.2.1.1). The index *i* also indicates the PMA lane number: 0, 1, 2, or 3.

On each transaction, tx_bit is assigned to C(*i*, *m*, *n*), where

i is the lane number
m is an index indicating the FEC codeword number and increments at the start of each codeword
n is an index indicating the bit number within a codeword

The start of a codeword is determined by the start parameter associated with the tx_bit parameter being equal to TRUE.

94.2.2.2 Overhead Frame

The PMA shall create a sequence of overhead frames by inserting 40 overhead bits for every 31280 FEC bits as specified in this section.

The FEC bits, C(*i*, *m*, *n*) are mapped into a continuous sequence of overhead frames. The overhead frame is 31320 bits in length.

Each bit in the overhead frame is denoted F(*i*, *p*, *q*), where:

i is the lane number
p is an index that indicates the frame number and increments at the start of each frame
q is an index that indicates the bit number within a frame with a range 1 to 31320

The first 40 bits of the frame, F(*i*, *p*, 1) to F(*i*, *p*, 40) are the overhead bits (see 94.2.2.3). The next 31280 bits, F(*i*, *p*, 41) to F(*i*, *p*, 31320) are composed of the bits from 23 consecutive FEC codewords.

The overhead bits are inserted in the frame as follows:

F(*i*, *p*, 1) = H(*i*, *p*, 1)
F(*i*, *p*, 2) = H(*i*, *p*, 2)
F(*i*, *p*, ...) = H(*i*, *p*, ...)
F(*i*, *p*, 40) = H(*i*, *p*, 40)

The FEC codeword bits are aligned such that F(*i*, *p*, 41) is the first bit of a codeword, e.g., F(*i*, *p*, 41) = C(*i*, *m*, 1). The FEC bits are inserted into the frame in the order in which they were received from the FEC, e.g., F(*i*, *p*, 42) = C(*i*, *m*, 2), F(*i*, *p*, 43) = C(*i*, *m*, 3), and so on. The method for aligning the FEC codeword with the start of the PMA frame is outside the scope of this standard.

94.2.2.3 Overhead

Editor's note (to be removed prior to final publication):

A proposal for overhead usage and content is required.

The overhead bits are denoted H(*i*, *p*, *k*), where

i is the lane number
p is an index that indicates the frame number and increments at the start of each frame
k is an index that indicates the header bit number with a range 1 to 40.

94.2.2.4 Termination Blocks

The PMA shall create a sequence of termination blocks by inserting a termination bit for every 45 overhead frame bits as specified in this section.

The termination block is 46 bits in length.

Each bit in a termination block is denoted T(*i*, *r*, *s*), where:

i is the lane number
r is an index indicating block number and increments at the start of each block
s is an index indicating the bit number within a termination block with a range 1 to 46

The first 45 bits of each termination block, T(*i*, *r*, 1) to T(*i*, *r*, 45), are overhead frame bits (see 94.2.2.2). The frame bits are aligned with the termination blocks such that the first bit of an overhead bit, F(*i*, *p*, 1), corresponds to the first bit of a termination block, T(*i*, *r*, 1). The 46th bit in each termination block,

T(*i*, *r*, 46), is set to zero. Overhead frame bits are mapped to the termination blocks in order of location within the overhead frame, e.g., T(*i*, *r*, 2) = F(*i*, *p*, 2), T(*i*, *r*, 3) = F(*i*, *p*, 3), and so on.

94.2.2.5 Gray Mapping

The PMA shall map consecutive pairs of bits to one of four Gray-coded symbols as specified in this section.

Each pair of bits, {A, B}, of each termination block are converted to a Gray-coded symbol with one of the four Gray-coded levels as follows:

{0, 0} maps to 0,
{0, 1} maps to 1,
{1, 1} maps to 2, and
{1, 0} maps to 3.

Gray-coded symbols corresponding to each termination block are denoted G(*i*, *r*, *t*), where:

i is the lane number
r is an index indicating the termination block number
t is an index indicating the symbol number within a termination block with a range 1 to 23.

Pairing of bits is such that the first two bits of each termination block, T(*i*, *r*, 1) and T(*i*, *r*, 2), form a pair. Each bit pair {T(*i*, *r*, 2*t*–1), T(*i*, *r*, 2*t*)} maps to {A, B} and the Gray-coded result is assigned to G(*i*, *r*, t). The gray-coded symbol G(*i*, *r*, 23) is formed from the last two bits of a termination block including one overhead frame bit and one termination bit (of value zero); thus G(*i*, *r*, 23), the Gray-coded termination symbol, always takes the value 0 or 3.

94.2.2.6 Precoding

The PMA shall precode the Gray-coded symbols as specified in this section.

The precoder output symbols are denoted, P(*i*, *r*, *t*), where:

i is the lane number
r is an index indicating the termination block number
t is an index indicating the symbol number within a termination block with a range 1 to 23.

For each Gray-coded symbol G(*i*, *r*, *t*), a precoded symbol, P(*i*, *r*, *t*) is determined by the following algorithm:

If $t = 23$ then
    $P(i, r, t) = G(i, r, t)$
Else If $t = 1$ then
    $P(i, r, t) = (G(i, r, t) - P(i, r-1, 23)) \bmod 4$
Else
    $P(i, r, ,t) = (G(i, r, t) - P(i, r, t-1)) \bmod 4$
End If

The Gray-coded termination symbol, G(*i*, *r*, 23), is always equal to either 0 or 3. The precoding algorithm above applies this symbol directly to the output, thus re-initializing the loop and ensuring that the precoded termination symbol,P(*i*, *r*, 23), is always either 0 or 3.

94.2.2.7 PAM4 encoding

The PMA shall encode each precoder output symbol to one of four PAM4 levels as specified in this section.

The PAM4 encoded symbols are denoted Q(*i*, *u*), where

i is the lane number

u is an index indicating the symbol number.

Each consecutive precoder output symbol, P(*i*, *r*, ,*t*), is mapped to one of four PAM4 levels and assigned to the PAM4 encoder output Q(*i*, 32*r*+*t*).

Mapping from the precoder output symbol P,(*i*, *r*, ,*t*) to a PAM4 encoded symbol Q(*i*, *u*) is as follows:

0 maps to –1,
1 maps to –1/3,
2 maps to +1/3, and
3 maps to +1.

The termination symbols after PAM4 encoding, Q(i, 32*r*+32), are always either –1 or +1.

94.2.2.8 PMD Interface

The PMA shall transmit each PAM4 encoded symbol, Q(*i*, *u*) to the PMD via the PMD:IS_UNITDATA_*i*(tx_symbol) primitive at a symbol transfer rate of 13.59375 GBd.

94.2.3 PMA Transmit EEE Operation

Editor's note (to be removed prior to final publication):

EEE operation for the 100GBASE-KP4 PHY has been adopted per barrass_01_0312 but details regarding EEE operation are not provided. A proposal on this topic is required.

94.2.4 PMA Receive Functional Specifications

The receive process shall recover the data encoded by the transmit process with a bit error ratio no higher than TBD after the overhead and termination bits have been removed. The process by which the receiver recovers the data to meet this requirement is outside the scope of this standard. The signal structure encoded by the transmitter process including the overhead bits, gray coding, and termination symbols may be leveraged by the receiver implementation at the discretion of the implementer. The remainder of this subclause specifies the receiver processes to reverse the transmitter encoding and report status.

In the received direction, the role of the 100GBASE-KP4 PMA is to adapt the PAM4 encoded signal from the PMD to a FEC encoded signal to be passed to the FEC for further processing. The adaptation processes shown in Figure 94–3 include PAM4 decoding, (1+D) mod 4 decoding, inverse Gray coding, remove termination bits, and remove overhead.

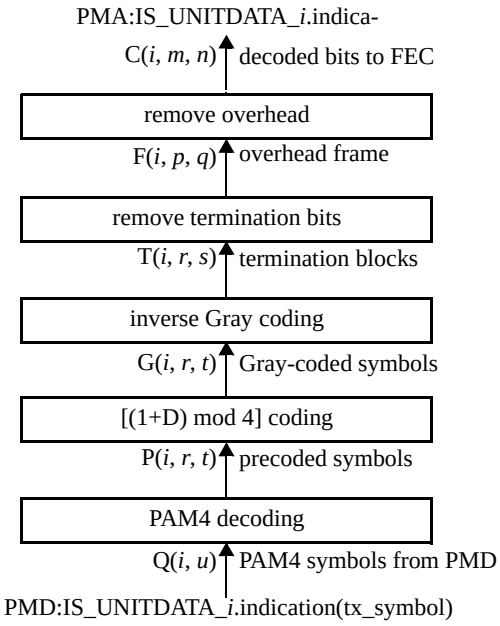


Figure 94–3—Receive adaptation process diagram

94.2.5 PMA Receive EEE Operation

Editor's note (to be removed prior to final publication):

The EEE baseline presentation, barrass_01_0312, does not provide functional details regarding the EEE for the 100GBASE-KP4 PMA. A proposal on this topic is required.

94.2.6 Skew constraints

Editor's note (to be removed prior to final publication):

The 100GBASE-KP4 PHY requires the FEC sublayer, which re-aligns all of the PCS lanes in the transmit direction and decodes aligned PMA lanes in the receive direction. Since the optional physical instantiation of the CAUI between the FEC sublayer and 100GBASE-KP4 PMA is not recommended, these requirements should limit the skew presented to the FEC sublayer from the PMA receive function. The skew is the combination of the skew contributed by the PMA, PMD, and the underlying medium.

A proposal on this topic is required.

94.2.7 Delay constraints

Delay considerations for the 100GBASE-KP4 PMA, PMD, AN, and medium are specified in 94.3.3.

94.2.8 Link status

The PMA shall provide link status information to the PMA client using the PMA:IS_SIGNAL.indication primitive (see 94.2.1.3). The PMA continuously monitors the link status reported by the PMD from the PMD:IS_SIGNAL.indication primitive, and uses this as input to Signal Indication Logic (SIL) to determine the link status to report to the PMA client. Other inputs to the SIL may include status of clock and data recovery on the lanes from the PMD and frame synchronization.

94.2.9 PMA local loopback mode

PMA local loopback shall be provided. This function involves looping back each input lane from the PMA service interface to the corresponding output lane on the PMA service interface. Each received instance of the PMA:IS_UNITDATA_*i*.request(tx_bit,start) primitive is looped back in the direction of the PMA client using the PMA:IS_UNITDATA_*i*.indication(rx_bit,start) primitive.

During local loopback, the PMA performs normal framing and precoding onto the lanes in the Tx direction toward the PMD service interface.

Ability to perform this function is indicated by the Local_loopback_ability status variable. If a Clause 45 MDIO is implemented, this variable is accessible through bit 1.8.0 (45.2.1.7.15). A device is placed in local loopback mode when the Local_loopback_enable control variable is set to one, and removed from local loopback mode when this variable is set to zero. If a Clause 45 MDIO is implemented, this variable is accessible through PMA/PMD control 1 register (bit 1.0.0, see 45.2.1.1.5).

94.2.10 PMA remote loopback mode (optional)

PMA remote loopback mode is optional. If implemented, it shall be as described in this subclause.

Remote loopback, if provided, should be implemented close enough to the PMD to maintain the bit sequence on each individual PMD lane. When remote loopback is enabled, each bit received over a lane of the service interface below the PMA via PMD:IS_UNITDATA_*i*.indication is looped back to the corresponding output lane toward the PMD via PMD:IS_UNITDATA_*i*.request.

During remote loopback, the PMA performs normal bit processing operation in the Rx direction towards the PMA client.

The ability to perform this function is indicated by the Remote_loopback_ability status variable. If a Clause 45 MDIO is implemented, this variable is accessible through bit 1.13.15 (45.2.1.12.1). A device is placed in remote loopback mode when the Remote_loopback_enable control variable is set to one, and removed from remote loopback mode when this variable is set to zero. If a Clause 45 MDIO is implemented, this variable is accessible through PMA/PMD Control register 1 (bit 1.0.1, see 45.2.1.1.4).

94.2.11 PMA test patterns

Editor's note (to be removed prior to final publication):

Test patterns as defined in Clause 83 are not directly applicable since they are defined for an PAM2 signal. A proposal on this topic is required.

94.2.12 PMA MDIO function mapping

Clause 45 specifies the optional MDIO capability that describes several variables that provide control and status information for and about the PMA. 45.2.1 describes the Management Data Input/Output (MDIO) Manageable Device (MMD) addresses".

94.3 Physical Medium Dependent (PMD) Sublayer

94.3.1 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 100GBASE-KP4 PMD. The service interface for this PMD is described in an abstract manner and does not imply any particular implementation. The PMD service interface supports the exchange of encoded data. The PMD translates the encoded data to and from signals suitable for the medium.

The PMD service interface based on the inter-sublayer service interface defined in 80.3. The PMD service interface primitives are summarized as follows:

PMD:IS_UNITDATA_*i*.request
PMD:IS_UNITDATA_*i*.indication
PMD:IS_SIGNAL.indication

If the optional EEE capability is supported, then the PMD service interface includes two additional primitives as follows:

PMD:IS_TX_MODE.request
PMD:IS_RX_MODE.indicate

Editor's note (to be removed prior to final publication):

Per barrass_01_0312, PMD transmit function behavior changes as a function of the tx_mode parameter.

DATA, SLEEP, or WAKE: normal behavior
ALERT: send the ALERT signal
FW: send a PMA-specific pattern that is TBD (PMD normal behavior?)
QUIET: disable the transmitter

Per barrass_01_0312, the PMD receive function infers rx_mode from the received signal.

Normal or rapid alignment markers: DATA, SLEEP, or WAKE (mapping not defined)
No signal: QUIET
Alert signal: ALERT
PMA-specific pattern for FW that is TBD: FW (PMA function?)

Given the ambiguity surrounding PMD behavior and the TBD signaling for ALERT, FW, etc., these functions have not been added to the clause. The functions will be added to the clause upon adoption of a detailed proposal describing the desired behavior.

94.3.1.1 PMD:IS_UNITDATA_i.request

The PMD:IS_UNITDATA_*i*.request (where *i*=0 to 3) primitive is used to define the transfer of four streams of data units from the PMA to the PMD.

94.3.1.1.1 Semantics of the service primitive

PMD:IS_UNITDATA_0.request(tx_symbol)
PMD:IS_UNITDATA_1.request(tx_symbol)
PMD:IS_UNITDATA_2.request(tx_symbol)
PMD:IS_UNITDATA_3.request(tx_symbol)

The data conveyed by PMD:IS_UNITDATA_i.request consists of four parallel continuous streams of encoded symbols, tx_symbol, one stream for each lane. Each of the tx_symbol parameters can take one of four values: –1, –1/3, +1/3, or +1.

94.3.1.1.2 When generated

The PMA continuously sends four parallel symbol streams PMD:IS_UNITDATA_*i*.request(tx_symbol) to the PMD, each at a nominal signaling rate of 13.59375 GBd.

94.3.1.1.3 Effect of receipt

Upon receiving each instance of PMD:IS_UNITDATA_*i*.request, the tx_symbol parameter is passed to the PMD transmit process corresponding to each stream.

94.3.1.2 PMD:IS_UNITDATA_i.indication

The PMD:IS_UNITDATA_*i*.indication (where *i*=0 to 3) primitive is used to define the transfer of four streams of data units from the PMD to the PMA.

94.3.1.2.1 Semantics of the service primitive

PMD:IS_UNITDATA_0.indication(rx_symbol)
PMD:IS_UNITDATA_1.indication(rx_symbol)
PMD:IS_UNITDATA_2.indication(rx_symbol)
PMD:IS_UNITDATA_3.indication(rx_symbol)

The data conveyed by PMD:IS_UNITDATA_0.indication to PMD:IS_UNITDATA_3.indication consists of 4 parallel continuous streams of encoded symbols, one stream for each lane. Each of the rx_symbol parameters can take one of four values: –1, –1/3, +1/3, or +1. The start parameter is TRUE to indicate that the concurrent rx_bit is the first bit of the first, second, third, or fourth FEC symbol in the FEC codeword and is otherwise FALSE.

94.3.1.2.2 When generated

The PMD continuously sends four parallel bit streams PMD:IS_UNITDATA_*i*.indication(tx_symbol) to the PMD client, each at a nominal signaling rate of 13.59375 GBd.

94.3.1.2.3 Effect of receipt

The effect of receipt of this primitive is undefined by the PMD.

94.3.1.3 PMD:IS_SIGNAL.indication

The PMD:IS_SIGNAL.indication primitive is generated by the PMD to the PMA to indicate the status of the PMD receive process. This primitive is generated by the PMD receive process to propagate the detection of severe error conditions (e.g., loss of synchronization) to the PMA.

94.3.1.3.1 Semantics of the service primitive

PMD:IS_SIGNAL.indication(SIGNAL_OK).

SIGNAL_DETECT in 100GBASE-KP4 indicates the successful completion of the start-up protocol on all four lanes. The SIGNAL_DETECT parameter can take on one of two values: OK or FAIL. When SIGNAL_DETECT = FAIL, the PMD:IS_UNITDATA_*i*.indication parameters are undefined. The SIGNAL_DETECT parameter maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

94.3.1.3.2 When generated

The PMD generates the PMD:IS_SIGNAL.indication primitive to the PMD client whenever there is change in the value of the SIGNAL_DETECT parameter.

94.3.1.3.3 Effect of receipt

The effect of receipt of this primitive is undefined by the PMD.

94.3.2 PCS requirements for Auto-Negotiation (AN) service interface

The PCS associated with this PMD is required to support the AN service interface primitive AN_LINK.indication defined in 73.9. (See 82.6.)

94.3.3 Delay constraints

The sum of the transmit and the receive delays contributed by the 100GBASE-KP4 PMA, PMD, AN, and the medium in one direction shall be no more than TBD bit times (TBD pause_quanta or TBD ns). It is assumed that the one way delay through the medium is no more than TBD bit times (TBD ns).

A description of overall system delay constraints and the definitions for bit times and pause_quanta can be found in 80.4 and its references.

94.3.4 Skew constraints

Editor's note (to be removed prior to final publication):

The 100GBASE-KP4 PHY requires the FEC sublayer, which re-aligns all of the PCS lanes in the transmit direction and decodes aligned PMA lanes in the receive direction. Since the optional physical instantiation of the CAUI between the FEC sublayer and 100GBASE-KP4 PMA is not recommended, these requirements should limit the skew presented to the FEC sublayer from the PMA receive function. The skew is the combination of the skew contributed by the PMA, PMD, and the underlying medium.

A proposal on this topic is required.

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. 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 80.5 and specified at the points SP1 to SP6 shown in Figure 80–4 and Figure 80–5.

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 TBD ns and the Skew Variation at SP2 is limited to TBD ps.

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

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

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 TBD ns and the Skew Variation at SP5 shall be less than TBD ns.

For more information on Skew and Skew Variation see 80.5.

94.3.5 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that provide control and status information for and about the PMD. If MDIO is implemented, it shall map MDIO control variables to PMD control variables as shown in Table 94–2, and MDIO status variables to PMD status variables as shown in Table 94–3.

Table 94–2—100GBASE-KP4 MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable	1.9.0	Global_PMD_transmit_disable
PMD transmit disable 3 to PMD transmit disable 0	PMD transmit disable	1.9.4 to 1.9.1	PMD_transmit_disable_3 to PMD_transmit_disable_0
Restart training	BASE-R PMD control	1.150.0	mr_restart_training
Training enable	BASE-R PMD control	1.150.1	mr_training_enable

Table 94–3—100GBASE-KP4 MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMD status variable
Fault	PMA/PMD status 1	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect	1.10.0	Global_PMD_signal_detect
PMD receive signal detect 3 to PMD receive signal detect 0	PMD receive signal detect	1.10.4 to 1.10.1	PMD_signal_detect_9 to PMD_signal_detect_0
Receiver status 3	BASE-R PMD status	1.151.12	rx_trained_3
Frame lock 3	BASE-R PMD status	1.151.13	frame_lock_3
Start-up protocol status 3	BASE-R PMD status	1.151.14	training_3
Training failure 3	BASE-R PMD status	1.151.15	training_failure_3
Receiver status 2	BASE-R PMD status	1.151.8	rx_trained_2
Frame lock 2	BASE-R PMD status	1.151.9	frame_lock_2
Start-up protocol status 2	BASE-R PMD status	1.151.10	training_2
Training failure 2	BASE-R PMD status	1.151.11	training_failure_2
Receiver status 1	BASE-R PMD status	1.151.4	rx_trained_1
Frame lock 1	BASE-R PMD status	1.151.5	frame_lock_1
Start-up protocol status 1	BASE-R PMD status	1.151.6	training_1
Training failure 1	BASE-R PMD status	1.151.7	training_failure_1
Receiver status 0	BASE-R PMD status	1.151.0	rx_trained_0
Frame lock 0	BASE-R PMD status	1.151.1	frame_lock_0
Start-up protocol status 0	BASE-R PMD status	1.151.2	training_0
Training failure 0	BASE-R PMD status	1.151.3	training_failure_0

94.3.6 PMD functional specifications

94.3.6.1 Link block diagram

One direction of a 100GBASE-KP4 link is shown in Figure 94–4.

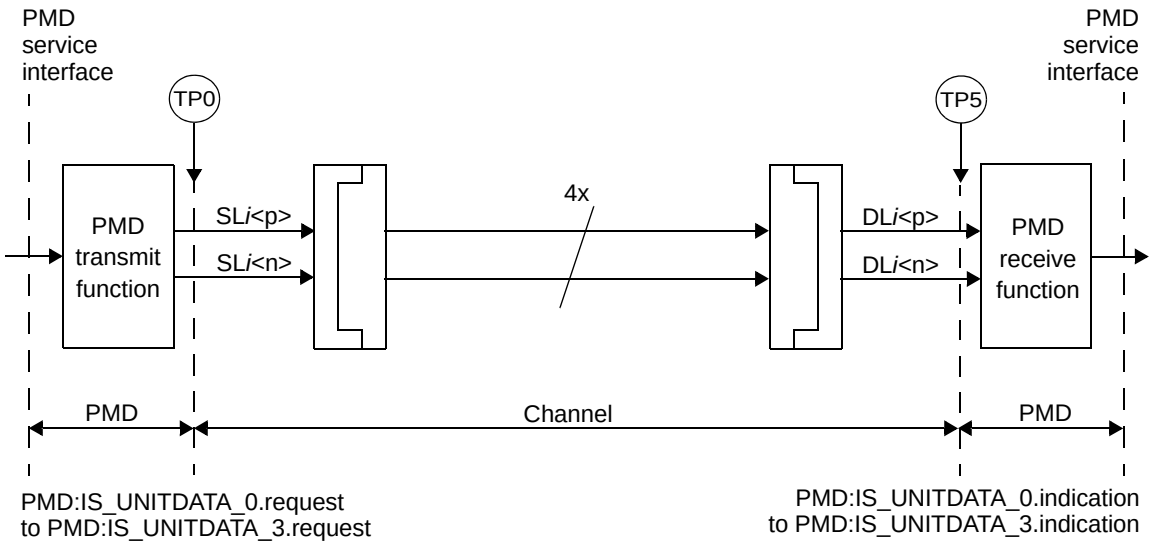


Figure 94–4—100GBASE-KP4 link (one direction is illustrated)

94.3.6.2 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 electrical streams. The four electrical signal streams shall then be delivered to the MDI, all according to the transmit electrical specifications in 94.3.11. A positive output voltage of SL*i*<p> minus SL*i*<n> (differential voltage) shall correspond to a positive tx_symbol value.

94.3.6.3 PMD Receive function

The PMD receive function shall convert the four electrical streams from the MDI 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. A positive input voltage of DL*i*<p> minus DL*i*<n> (differential voltage) shall correspond to a positive tx_symbol value.

94.3.6.4 Global PMD signal detect function

Editor's note (to be removed prior to final publication):

The reference to PMD:IS_SIGNAL.indication(SIGNAL_DETECT) must be reconciled with subclause 80.3 and 94.2.1.3.

The Global PMD signal detect function shall continuously report the message PMD:IS_SIGNAL.indication (SIGNAL_DETECT) to the PMD service interface. SIGNAL_DETECT, while normally intended to be an indicator of signal presence, is used by the 100GBASE-KP4 PMD to indicate the successful completion of

the start-up protocol on all lanes. The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

SIGNAL_DETECT shall be set to FAIL following system reset or the manual reset of the training state diagram. Upon successful completion of training on all lanes, SIGNAL_DETECT shall be set to OK.

If training is disabled by management, SIGNAL_DETECT shall be set to OK.

If the MDIO interface is implemented, then Global_PMD_signal_detect (1.10.0) shall be continuously set to the value of SIGNAL_DETECT as described in 45.2.1.9.7.

94.3.6.5 PMD lane-by-lane signal detect function

When the MDIO is implemented, each PMD_signal_detect_*i* value, where *i* represents the lane number in the range 0 to 3, shall be continuously updated as described in the following two paragraphs.

PMD_signal_detect_0 (1.10.1), PMD_signal_detect_1 (1.10.2), PMD_signal_detect_2 (1.10.3), and PMD_signal_detect_3 (1.10.4) shall be set to one or zero depending on whether a particular lane's signal_detect, as defined by the training state diagram in Figure 72–5, returns true or false.

94.3.6.6 Global PMD transmit disable function

The Global_PMD_transmit_disable function is optional. When implemented, it allows all of the transmitters to be disabled with a single variable.

a) When Global_PMD_transmit_disable variable is set to one, this function shall turn off all of the transmitters such that each transmitter drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage in Table 94–4.

b) If a PMD_fault (94.3.7) is detected, then the PMD may turn off the electrical transmitter in all lanes.

c) Loopback, as defined in 94.3.6.8, shall not be affected by Global_PMD_transmit_disable.

94.3.6.7 PMD lane-by-lane transmit disable function

The PMD_transmit_disable_*i* function (where *i* represents the lane number in the range 0:3) is optional and allows the electrical transmitter in each lane to be selectively disabled. When this function is supported, it shall meet the following requirements:

a) When a PMD_transmit_disable_*i* variable is set to one, this function shall turn off the transmitter associated with that variable such that it drives a constant level (i.e., no transitions) and does not exceed the maximum differential peak-to-peak output voltage specified in Table 94–4.

b) If a PMD_fault (94.3.7) is detected, then the PMD may set each PMD_transmit_disable_*i* to one, turning off the electrical transmitter in each lane.

c) Loopback, as defined in 94.3.6.8, shall not be affected by PMD_transmit_disable_*i*.

94.3.6.8 Loopback mode

Local loopback mode shall be provided by the adjacent PMA (see 83.5.8) as a test function to the device. When loopback mode is enabled, transmission requests passed to each transmitter are sent directly to the corresponding receiver, overriding any signal detected by each receiver on its attached link. Note that loopback mode does not affect the state of the transmitter, which continues to send data (unless disabled).

Control of the loopback function is specified in 45.2.1.1.5.

NOTE 1—The signal path that is exercised in the loopback mode is implementation specific, but it is recommended that this signal path encompass as much of the circuitry as is practical. The intention of providing this loopback mode of operation is to permit diagnostic or self-test functions to test the transmit and receive data paths using actual data. Other loopback signal paths may also be enabled independently using loopback controls within other devices or sublayers.

NOTE 2—Placing a network port into loopback mode can be disruptive to a network.

94.3.7 PMD_fault function

If the MDIO is implemented, PMD_fault is the logical OR of PMD_receive_fault, PMD_transmit_fault, and any other implementation specific fault.

94.3.8 PMD transmit fault function

The PMD_transmit_fault function is optional. The faults detected by this function are implementation specific, but should not include the assertion of the Global_PMD_transmit_disable function.

If a PMD_transmit_fault (optional) is detected, then the Global_PMD_transmit_disable function should also be asserted.

If the MDIO interface is implemented, then this function shall be mapped to the Transmit fault bit as specified in 45.2.1.7.4.

94.3.9 PMD receive fault function

The PMD_receive_fault function is optional. The faults detected by this function are implementation specific.

If the MDIO interface is implemented, then this function shall contribute to the Receive fault bit as specified in 45.2.1.7.5.

94.3.10 PMD control function

Editor's note (to be removed prior to final publication):

The baseline proposal does not specify a requirement for a training mechanism. However, a training mechanism is assumed and some of the concepts of this section are relevant to many of the transmitter measurements in the following subclauses. A proposal on this topic is required.

Each lane of the 100GBASE-KP4 PMD shall use the same control function as 10GBASE-KR, as defined in 72.6.10.

The random seed for the training pattern described in 72.6.10.2.6 shall be different for each of the lanes.

The variables rx_trained_*i*, frame_lock_*i*, training_*i*, and training_failure_*i* (where *i* goes from 0 to 3) report status for each lane and are equivalent to rx_trained, frame_lock, training, and training_failure as defined in 72.6.10.3.1.

If the MDIO interface is implemented, then this function shall map these variables to the appropriate bits in the BASE-R PMD status register (Register 1.151) as specified in 45.2.1.80.

94.3.11 PMD Transmitter electrical characteristics

Transmitter characteristics measured at TP0a are summarized in Table 94–4.

Table 94–4—Summary of transmitter characteristics at TP0a

Parameter	Subclause reference	Value	Units
Signaling rate	94.3.11.2	13.59375 ± 100 ppm	GBd
Differential peak-to-peak output voltage (max.) Transmitter disabled[a] Transmitter enabled	94.3.11.3	 30 1200	 mV mV
Common-mode DC output voltage (max.)	94.3.11.3	1.9	V
Common-mode DC output voltage (min.)	94.3.11.3	0	V
Common-mode AC output voltage (RMS, max.)	94.3.11.3	30	mV
Differential output return loss (min.)	94.3.11.4	Equation (94–3)	dB
Common-mode output return loss (min.)	94.3.11.4	Equation (94–4)	dB
Transition time (20-80%, min.), no equalization[b]	94.3.11.5	TBD	ps
Normalized distortion factor (max.)	94.3.11.6	0.06	
Output waveform Steady-state voltage v_f (max.) Steady-state voltage v_f (min.) Linear fit pulse peak (min.) Normalized RMS linear fit error (max.) Normalized coefficient step size (min.) Normalized coefficient step size (max.) Pre-cursor full-scale range (min.) Post-cursor full-scale range (min.)	94.3.11.7	 0.6 TBD TBD TBD TBD TBD TBD TBD	 V V V — — — — —
Far-end output noise (max.) Low insertion loss channel High insertion loss channel	94.3.11.8	 TBD TBD	 mV mV
Output jitter (max.) Effective random jitter Even-odd jitter Total jitter excluding data dependent jitter	94.3.11.9	 TBD TBD TBD	 UI UI UI

[a]The transmitter for lane *i* is disabled when either Global_PMD_transmit_disable or PMD_transmit_disable_*i* is set to one.
[b]Transmit equalization may be disabled by asserting the preset control defined in Table 45–60 and 45.2.1.81.3.

94.3.11.1 Test Fixture

The test fixture of Figure 94–5 or its equivalent, is required for measuring the transmitter specifications described in 94.3.11.

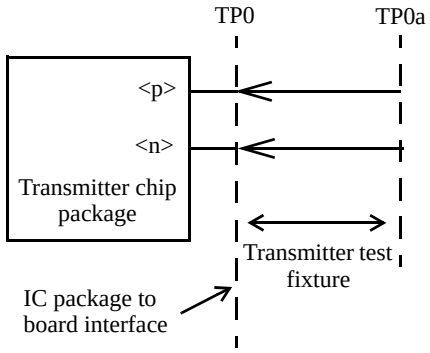


Figure 94–5—Transmitter test fixture and test points.

94.3.11.1.1 Test fixture impedance

The differential load impedance applied to the transmitter output by the test fixture depicted in Figure 94–5 shall be 100 Ω. The differential return loss, in dB with f in MHz, of the test fixture shall meet the requirements of Equation (94–1) and Equation (94–2).

TBD (94–1)

TBD (94–2)

94.3.11.2 Signaling rate and range

The 100GBASE-KP4 signaling rate shall be 13.59375 GBd ± 100 ppm per lane.

94.3.11.3 Signal levels

The differential output voltage v_{di} is defined to be SL*i*<p> minus SL*i*<n>. The common-mode output 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 94–6.

For a square wave test pattern with a 2 UI period, the peak-to-peak differential output voltage shall be less than or equal to 1200 mV regardless of the transmit equalizer setting. The peak-to-peak differential output voltage shall be less than or equal to 30 mV when the transmitter is disabled (refer to 94.3.6.6 and 94.3.6.7). The differential output voltage test pattern shall consist of no fewer than eight symbols of alternating polarity.

The common-mode DC output voltage shall be between 0 V and 1.9 V with respect to signal ground. The common-mode AC output voltage shall be less than or equal to 30 mV RMS with respect to signal ground.

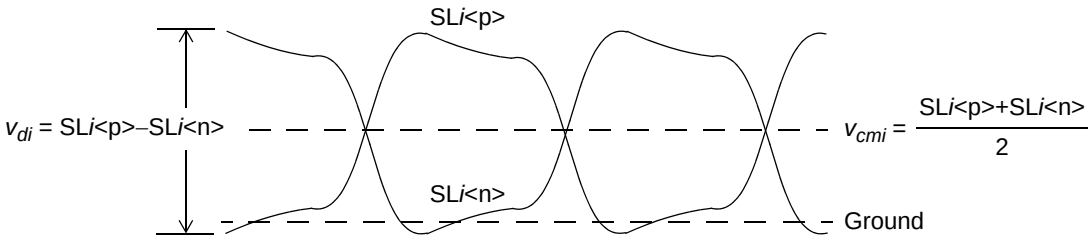


Figure 94–6—Transmitter output voltage definitions

94.3.11.4 Transmitter output return loss

The differential output return loss, in dB, of the transmitter shall meet Equation (94–3). This output impedance requirement applies to all valid output levels. The reference impedance for differential return loss measurements shall be 100 Ω.

TBD (94–3)

The common-mode output return loss, in dB, of the transmitter shall meet Equation (94–4). This output impedance requirement applies to all valid output levels. The reference impedance for common-mode return loss measurements shall be 25 Ω.

TBD (94–4)

94.3.11.5 Transition time

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 (see 68.6.2).

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 zeros. In this case, the 0% level and the 100% level may be estimated as ZeroLevel and ZeroLevel+MeasuredOMA in the TWDP code (see 68.6.6.2) or as the average signal within windows from –3 UI to –2 UI and from 2 UI to 3 UI relative to the edge.

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.

The transition times shall be greater than or equal to TBD ps when transmit equalization is disabled. Transmit equalization may be disabled by asserting the preset control defined in Table 45–60 and 45.2.1.81.3.

94.3.11.6 Transmitter linearity

Transmitter linearity is measured using a special 80-symbol pattern. The pattern is a sequence of 10 voltage levels each 8 unit intervals in duration. The 10 levels correspond to the following set of PAM4 symbols: {–1,–1/3,+1/3,+1,–1,+1,–1,+1,+1/3,–1/3}. The resulting waveform is shown in Figure 94–8. The transmitter is

configured with coefficients $c(-1)$ and $c(+1)$ set to zero.

Each transmitted level, V_A, V_B, V_C, and V_D, is measured after 3 to 7 unit intervals of settling. The normalized distortion factor for each level, D_A, D_B, D_C, and D_D, respectively, is calculated based on equations 94–5 to 94–11. The normalized distortion factor for of the four levels shall be less than 0.06.

$$V_{Low} = \frac{(V_C - V_B)}{2} \quad (94–5)$$

$$V_{High} = \frac{(V_D - V_A)}{6} \quad (94–6)$$

$$V_{Avg} = \frac{(V_{High} - V_{Low})}{2} \quad (94–7)$$

$$D_A = \frac{\left|V_{High} + \frac{V_A}{3}\right|}{V_{Avg}} \quad (94–8)$$

$$D_B = \frac{|V_{Low} + V_B|}{V_{Avg}} \quad (94–9)$$

$$D_C = \frac{|V_{Low} - V_C|}{V_{Avg}} \quad (94–10)$$

$$D_D = \frac{\left|V_{High} - \frac{V_D}{3}\right|}{V_{Avg}} \quad (94–11)$$

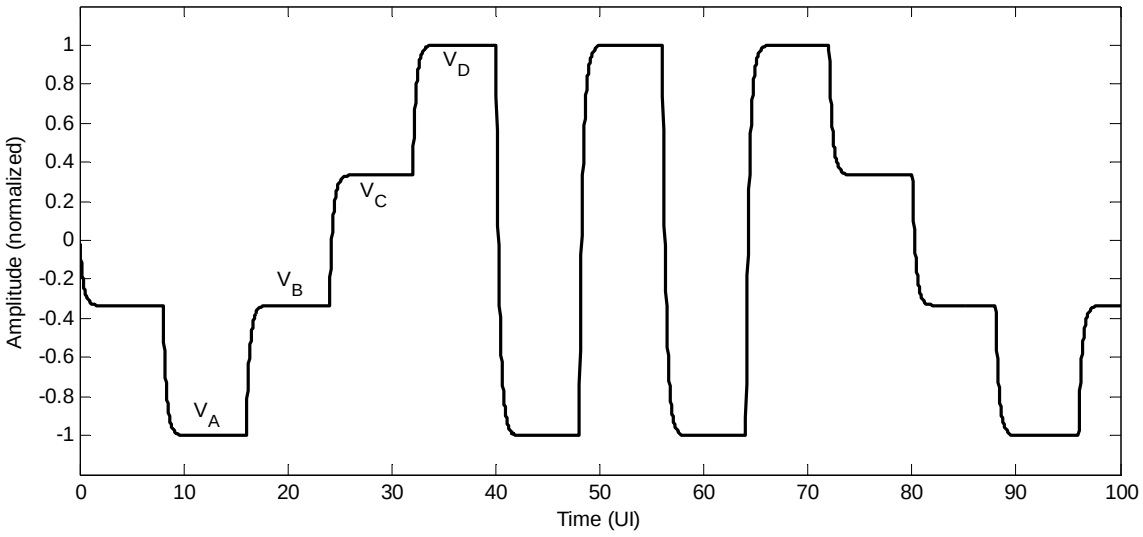


Figure 94–7—Transmitter linearity test pattern

94.3.11.7 Transmitter output waveform

The 100GBASE-KP4 transmit function includes programmable equalization to compensate for the frequency-dependent loss of the channel and facilitate data recovery at the receiver. The functional model for the transmit equalizer is the three tap transversal filter shown in Figure 94–8.

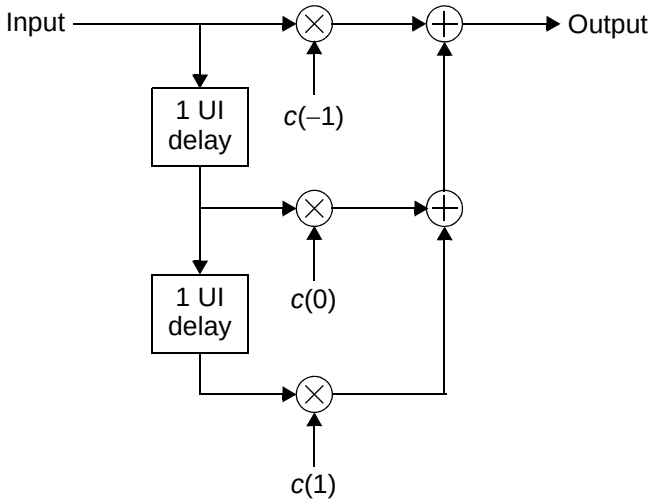


Figure 94–8—Transmit equalizer functional model

The state of the transmit equalizer and hence the transmitted output waveform may be manipulated via the PMD control function defined in 94.3.10 or via the management interface. The transmit function responds to a set of commands issued by the link partner's receive function and conveyed by a back-channel communications path.

This command set includes instructions to:

Increment coefficient $c(i)$.
Decrement coefficient $c(i)$.
Hold coefficient $c(i)$ at its current value.
Set the coefficients to a pre-defined value (preset or initialize).

In response, the transmit function relays status information to the link partner's receive function. The status messages indicate that:

The requested update to coefficient $c(i)$ has completed (updated).
Coefficient $c(i)$ is at its minimum value.
Coefficient $c(i)$ is at its maximum value.
Coefficient $c(i)$ is ready for the next update request (not_updated).

The transmitter output waveform is characterized using the procedure described in 85.8.3.3. The parameters of the linear pulse fit and equalizer are summarized in Table 94–5.

Table 94–5—Linear fit pulse and equalizer parameters

Description	Symbol	Value	Units
Linear fit pulse length	N_p	8	UI
Linear fit pulse delay	D_p	2	UI
Equalizer length	N_w	8	UI
Equalizer delay	D_w	2	UI

94.3.11.7.1 Steady state voltage and linear fit pulse peak

The steady state voltage and linear fit pulse peak values shall be verified after the transmit equalizer coefficients have been set to the "preset" values.

The steady state voltage v_f is defined to be the sum of the linear fit pulse $p(k)$ divided by M (refer to 85.8.3.3 step 3). The steady state voltage shall be greater than or equal to TBD V and less than or equal to 0.6 V.

The peak value of $p(k)$ shall be greater than TBD × v_f.

94.3.11.7.2 Linear fit error

For any configuration of the transmit equalizer, the RMS value of the error between the linear fit and the measured waveform, $e(k)$, normalized to the peak value of the linear fit pulse, $p(k)$, shall be less than or equal to TBD.

94.3.11.7.3 Coefficient initialization

When the PMD enters the INITIALIZE state of the Training state diagram (Figure 72–5) or receives a valid request to "initialize" from the link partner, the coefficients of the transmit equalizer shall be configured such that the ratio $(c(0)+c(1)-c(-1))/(c(0)+c(1)+c(-1))$ is TBD and the ratio $(c(0)-c(1)+c(-1))/(c(0)+c(1)+c(-1))$ is TBD. These requirements apply upon the assertion a coefficient status report of "updated" for all coefficients.

94.3.11.7.4 Coefficient step size

The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to "increment" that coefficient shall be between 0.0083 and 0.05. The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to "decrement" that coefficient shall be between –0.05 and –0.0083.

The change in the normalized amplitude of the coefficient is defined to be the difference in the value measured prior to the assertion of the "increment" or "decrement" request (e.g., the coefficient update request for all coefficients is "hold") and the value upon the assertion of a coefficient status report of "updated" for that coefficient.

94.3.11.7.5 Coefficient range

When sufficient “increment” or “decrement” requests have been received for a given coefficient, the coefficient will reach a lower or upper bound based on the coefficient range or restrictions placed on the minimum steady state differential output voltage or the maximum peak-to-peak differential output voltage.

With $c(-1)$ set to zero and both $c(0)$ and $c(1)$ having received sufficient “decrement” requests so that they are at their respective minimum values, the ratio $(c(0)-c(1))/(c(0)+c(1))$ shall be greater than or equal to 4.

With $c(1)$ set to zero and both $c(-1)$ and $c(0)$ having received sufficient “decrement” requests so that they are at their respective minimum values, the ratio $(c(0)-c(-1))/(c(0)+c(-1))$ shall be greater than or equal to 1.54.

Note that a coefficient may be set to zero by first asserting the preset control and then manipulating the other coefficients as required by the test.

94.3.11.8 Transmitter far-end output noise

The transmitter far-end output noise is a source of noise in addition to the channel integrated crosstalk noise (ICN) specified in 94.4.3. The transmitter far-end output noise is characterized using the procedure defined in 85.8.3.2.

The far-end transmitter output noise is characterized using two reference channels:

A “low-loss” channel with TBD dB insertion loss at 6.875 GHz and ICN that meets the requirements of 94.4.3.
A “high-loss” channel with TBD dB insertion loss at 6.875 GHz and ICN that meets the requirements of 94.4.3.

For the low-loss channel, the far-end transmitter output noise shall be less than or equal to TBD mV RMS. For the high-loss channel, the far-end transmitter output noise shall be less than or equal to TBD mV RMS.

94.3.11.9 Transmitter output jitter

Even-odd jitter is characterized using the procedure defined in 92.8.3.8. Even-odd jitter shall be less than or equal to TBD UI regardless of the transmit equalization setting.

Total jitter (TJ) is characterized using the procedure defined in 92.8.3.8. Data dependent jitter (DDJ) is characterized using the procedure defined in 85.8.3.8. The total jitter excluding data dependent jitter, shall be less than or equal to TBD UI regardless of the transmit equalization setting.

Editor’s note (to be removed prior to final publication):

Clause 85 was not clear on what was meant by “excluding” DDJ from TJ. In this draft, it has been interpreted to be the value of TJ as defined in 92.8.3.8 minus the value of DDJ as defined in 85.8.3.8.

The effective random jitter (RJ) is characterized using the procedure defined in 92.8.3.8. The effective random jitter shall be less than or equal to TBD UI regardless of the transmit equalization setting.

For the purpose of measuring TJ, DDJ, and RJ, the effect of a single-pole high-pass filter with a 3 dB frequency of TBD MHz is applied to the jitter.

Editor's note (to be removed prior to final publication):

Clause 72 specifies that transmit equalization is off (preset) for the measurement of jitter. Clause 85 does not specify the conditions for jitter measurement and this is interpreted to mean that it applies for all transmit equalization settings. There has been some discussion as to which approach is appropriate.

The test pattern for TJ, DDJ, and RJ measurements is TBD.

94.3.12 PMD Receiver electrical characteristics

Receiver characteristics measured at TP5a are summarized in Table 94–6.

Table 94–6—Summary of receiver characteristics at TP5a

Parameter	Subclause reference	Value	Units
Differential input return loss (min.)	94.3.12.2	Equation (94–14)	dB
Common-mode input return loss (min.)	94.3.12.2	Equation (94–15)	dB
Differential to common-mode return loss (min.)	94.3.12.2	Equation (94–16)	dB

94.3.12.1 Test Fixture

The test fixture of Figure 94–5 or its equivalent, is required for measuring the receiver specifications described in 94.3.12.

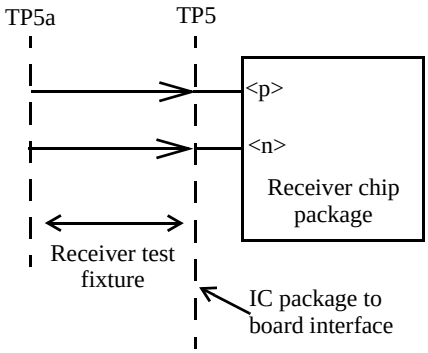


Figure 94–9—Receiver test fixture and test points.

94.3.12.1.1 Test fixture impedance

The differential load impedance applied to the receiver output by the test fixture depicted in Figure 94–5 shall be 100 Ω. The differential return loss, in dB with f in MHz, of the test fixture shall meet the requirements of Equation (94–12) and Equation (94–13).

TBD (94–12)

TBD (94–13)

94.3.12.2 Receiver input return loss

The differential input return loss, in dB, of the receiver shall meet Equation (94–14). The reference impedance for differential return loss measurements shall be 100 Ω.

TBD (94–14)

The common-mode input return loss, in dB, of the receiver shall meet Equation (94–15). The reference impedance for common-mode return loss measurements shall be 25 Ω

TBD (94–15)

The differential to common-mode return loss, in dB, of the receiver shall meet Equation (94–16).

TBD (94–16)

94.3.12.3 Receiver interference tolerance

Editor's note (to be removed prior to final publication):

Test 1 is intended to be on a low loss channel with ~15 dB IL at Nyquist, while Test 2 is intended to be a on a high loss channel with ~33 dB IL at Nyquist.

Receiver interference tolerance is characterized using the procedure defined in Annex 69A. The receiver shall satisfy the requirements for interference tolerance summarized in Table 94–7. FEC is not included for tests 1 and 2. FEC is included for tests 3 and 4.

Table 94–7—Receiver interference tolerance parameters

Parameter	Test 1 values	Test 2 values	Units
Maximum BER without FEC	3×10^{-4}	3×10^{-4}	
Maximum BER with FEC	10^{-12}	10^{-12}	
Channel insertion loss at 6.875 GHz Real part of a_0, min.[a] Real part of a_1, min. Real part of a_2, min. Real part of a_4, min.	TBD TBD TBD TBD TBD	TBD TBD TBD TBD TBD	dB $Hz^{-1/2}$ Hz^{-1} Hz^{-2}
Channel noise excluding TX-RX re-reflection noise	0	0	mV
Applied peak-to-peak Sinusoidal Jitter[b]	TBD	TBD	UI
Applied peak-to-peak Random Jitter[c]	TBD	TBD	UI
Applied even-odd jitter	TBD	TBD	UI
Applied RMS broadband noise	TBD	TBD	mV

[a]For each test channel, a_0 is limited to a maximum value of 1.5 and a_1, a_2, and a_4 are limited to a maximum value of 0.

[b]The frequency of the sinusoid must be greater than 100 MHz.
[c]Random Jitter is specified at a BER of 10^{-12}.

94.3.13 AC-coupling

The 100GBASE-KP4 transmitter shall be AC-coupled to the receiver. Common-mode specifications are defined as if the DC-blocking capacitor is implemented between TP0 and TP5. Should the capacitor be implemented outside TP0 and TP5, the common-mode specifications in Table 94–4 may not be appropriate.

The impact of a DC-blocking capacitor implemented between TP0 and TP5 is accounted for within the channel specifications. Should the capacitor be implemented outside TP0 and TP5, it is the responsibility of implementers to consider any necessary modifications to common-mode and channel specifications required for interoperability as well as any impact on the verification of transmitter and receiver compliance.

Editor's note (to be removed prior to final publication):

The following requirement is against the channel not the receiver. Consider moving this requirement to the channel characteristics section.

The low frequency 3 dB cutoff of the AC coupling shall be less than 50 kHz.

94.4 Channel characteristics

94.4.1 Channel operating margin

The channel operating margin (COM) computed using the procedure in Annex 93A and the parameters in Table 94–8 shall be greater than or equal to TBD dB.

94.4.2 Channel insertion loss

The insertion loss, in dB, of the channel is recommended to meet Equation (94–17). The insertion loss limit is shown Figure 94–10.

$$IL(f) \le IL_{max}(f) = \left\{ \begin{array}{ll} A_{max}(f) + 0.8 + 2.0 \times 10^{-10} \cdot f & f_{min} \le f \le f_2 \\ A_{max}(f_2) + 0.8 + 2.0 \times 10^{-10} \cdot f_2 + 1 \times 10^{-8} \cdot (f - f_2) & f_2 < f \le f_{max} \end{array} \right\} \quad (94–17)$$

where

$IL(f)$ is the insertion loss at frequency f
$IL_{max}(f)$ is the maximum allowable insertion loss at frequency f
A_{max} is given in Equation (94–18)
f is the measurement frequency in Hz
f_{min} = 0.05 GHz
f_2 = 7.0 GHz
f_{max} = 15 GHz

Table 94–8—Channel operating margin parameters

Parameter	Symbol	Value	Units
Signaling rate	f_b	13.59375	GHz
Maximum start frequency	f_{min}	0.05	GHz
Maximum frequency step	Δf	0.01	GHz
Transmitter reflection coefficient DC value Reference frequency	Γ_1 Γ_{01} f_1	 TBD TBD	 V/V GHz
Receiver reflection coefficient DC value Reference frequency	Γ_2 Γ_{02} f_2	 TBD TBD	 V/V GHz
Transmitter differential peak output voltage Victim Far-end aggressor Near-end aggressor	 A_v A_f A_n	 0.4 0.4 0.6	 V V V
Transmitter 3 dB bandwidth Victim Far-end aggressor Near-end aggressor	 f_v f_f f_n	 $0.55 \times f_b$ $0.55 \times f_b$ f_b	 GHz GHz GHz
Receiver 3 dB bandwidth	f_r	$0.75 \times f_b$	
Transmitter equalizer, pre-cursor coefficient Minimum value Maximum value Step size	$c(-1)$	 –0.10 0.00 0.02	 — — —
Transmitter equalizer, post-cursor coefficient Minimum value Maximum value Step size	$c(1)$	 –0.40 0.00 0.02	 — — —
Continuous time filter, DC gain Minimum value Maximum value Step size	g_{DC}	 –12 0 1	 dB dB dB
Number of signal levels	L	4	—
Number of samples per unit interval	M	32	—
Victim single bit response exception window	W	TBD	UI
Normalized RMS Gaussian noise	σ_G	0.01	—
Normalized peak dual-Dirac noise	A_{DD}	0.1	—
Target uncorrected symbol error ratio	SER_0	10^{-5}	—
Minimum channel operating margin	COM_0	TBD	dB

$$A_{max}(f) = 20 \cdot \log_{10}(e) \cdot (b_1 \cdot \sqrt{f} + b_2 \cdot f + b_3 \cdot f^2 + b_4 \cdot f^3) \quad (94–18)$$

where

$b_1 = 2.00 \times 10^{-5}$
$b_2 = 1.10 \times 10^{-10}$

$b_3 = 3.20 \times 10^{-20}$
$b_4 = -1.20 \times 10^{-30}$

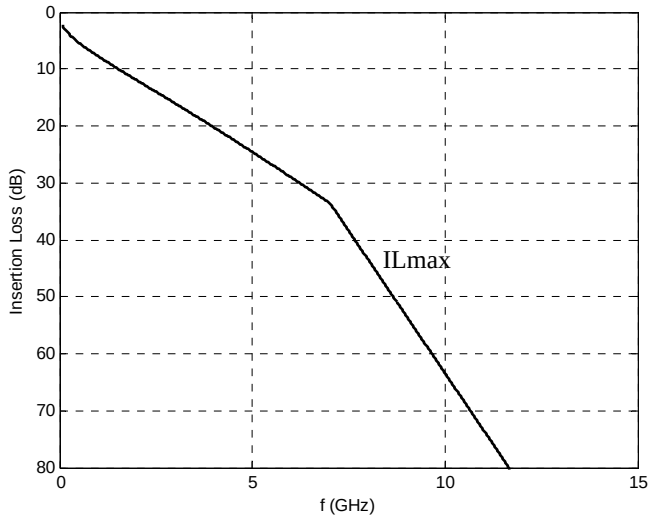


Figure 94–10—Channel Insertion Loss limit

94.4.3 Channel Integrated Crosstalk Noise (ICN)

Editor's note (to be removed prior to final publication):

This is a placeholder to be referenced by the transmitter far-end output noise specification (refer to 94.3.11.8).

94.5 Environmental specifications

94.5.1 General safety

All equipment subject to this clause shall conform to applicable sections (including isolation requirements) of IEC 60950-1.

94.5.2 Network safety

The designer is urged to consult the relevant local, national, and international safety regulations to ensure compliance with the appropriate requirements.

94.5.3 Installation and maintenance guidelines

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

94.5.4 Electromagnetic compatibility

A system integrating the 100GBASE-KP4 PHY shall comply with applicable local and national codes for the limitation of electromagnetic interference.

94.5.5 Temperature and humidity

A system integrating the 100GBASE-KP4 PHY is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration). Specific requirements and values for these parameters are considered to be beyond the scope of this standard.

94.6 Protocol implementation conformance statement (PICS) proforma for Clause 94, Physical Medium Attachment (PMA) and Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KP4[12]

94.6.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 93, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KP4, shall complete the following protocol implementation conformance statement (PICS) proforma.

A detailed description of the symbols used in the PICS proforma, along with instructions for completing the PICS proforma, can be found in Clause 21.

94.6.2 Identification

94.6.2.1 Implementation identification

Supplier[1]	
Contact point for enquiries about the PICS[1]	
Implementation Name(s) and Version(s)[1,3]	
Other information necessary for full identification—e.g., name(s) and version(s) for machines and/or operating systems; System Name(s)[2]	
NOTE 1— Required for all implementations. NOTE 2— May be completed as appropriate in meeting the requirements for the identification. NOTE 3—The terms Name and Version should be interpreted appropriately to correspond with a supplier's terminology (e.g., Type, Series, Model).	

94.6.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3bj-20XX, Clause 93, Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KP4
Identification of amendments and corrigenda to this PICS proforma that have been completed as part of this PICS	
Have any Exception items been required? No [ ] Yes [ ] (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3bj-20XX.)	

Date of Statement	

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

94.6.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
CGMII	CGMII	94.1	Interface is supported	O	Yes [ ] No [ ]
PCS	100GBASE-R PCS	94.1		M	Yes [ ]
FEC	Forward error correction	94.1		M	Yes [ ] No [ ]
PMA	100GBASE-R PMA	94.1		M	Yes [ ]
CAUI	CAUI	94.1	Interface is supported	O	Yes [ ] No [ ]
AN	Auto-negotiation	94.1		M	Yes [ ]
DC	Delay constraints	94.3.3		M	Yes [ ]
DSC	Skew constraints	94.3.4		M	Yes [ ]
*MD	MDIO capability	94.3.5	Registers and interface supported	O	Yes [ ] No [ ]

94.6.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and baseband medium, type 100GBASE-KP4

Editor's note (to be removed prior to final publication):

PICS proforma tables will be generated when the content of the clause stabilizes.

Annex 83A

(normative)

40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)

Editor's note (to be removed prior to final publication):

The IEEE P802.3bj Task Force and IEEE 802.3 Working Group approved changes to the project objectives, 5 criteria responses, and PAR that increased the scope of the project to add the optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10. However, the change of scope is still pending LMSC EC and SASB approval. It is expected that the EC will consider the change at the November 2012 meeting and, based on that, the SASB will consider the change at their December 2012 meeting.

Based on these completed and pending actions, this clause is changed in a manner that is consistent with the increased scope. Should the PAR modification not be approved, any changes pertaining to an optional Energy Efficient Ethernet capability for 40GBASE-CR4, 40GBASE-KR4, and 100GBASE-CR10 will be removed.

Add the following after 83A.3.2:

83A.3.2a EEE operation

If the optional Energy Efficient Ethernet (EEE) capability is supported (see Clause 78, 78.3) then the inter-sublayer service interface includes two additional primitives as described in 83.3 and may also support CAUI shutdown.

Editor's note (to be removed prior to final publication):

The following behavior has been proposed but not adopted by the Task Force for EEE behavior:

In the ingress direction, the CAUI shall transmit a repeating 16-bit pattern, hexadecimal 0xFF00 while parameter rx_mode = ALERT.

In the ingress direction, in addition to the transmit disable function defined by 83.5.12 and 83.5.12.1 (references changed by another comment), the CAUI shall transmit the PRBS31 pattern defined in 83.5.10 when rx_mode = QUIET. The requirement to disable the transmitters takes precedance over the PRBS test pattern transmission.

In the ingress direction, a PMA that is connected by a CAUI to a separated PMA may infer the state of rx_mode by observing this behavior of the CAUI signals.

Add the following at the end of 83A.3.3.1:

83A.3.3.1.1 Amplitude and swing

For EEE capability, the CAUI transmitter lane's differential peak-to-peak output voltage shall be less than TBD mV within TBD ns of tx_mode changing to QUIET in the egress direction or rx_mode changing to

QUIET in the ingress direction. Furthermore, the CAUI transmitter lane's differential peak-to-peak output voltage shall be greater than TBD mV within TBD ns of tx_mode ceasing to be QUIET in the egress direction or rx_mode ceasing to be QUIET in the ingress direction.

Editor's note (to be removed prior to final publication):

The baseline highlights functions that are still to be defined:

The state of tx_mode and rx_mode must both be communicated across CAUI.
When tx_mode is in state ALERT, the definition of the ALERT signaling must be defined.
When tx_mode in in state FW, some signaling must be defined.
The receiver must infer the state of rx_mode from the received signal (for states DATA, ALERT & FW)

Add the following after 83A.3.3.5 for the CAUI shutdown:

83A.3.3.6 Global transmit disable function

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 may turn off all the transmitter lanes for the egress direction CAUI if PEASE is asserted and tx_mode is QUIET. In the ingress direction, the PMA may turn off all the transmitter lanes for the ingress direction CAUI if PIASE is asserted and rx_mode is QUIET. In both directions, the transmit disable function shall turn on all transmitter lanes after tx_mode or rx_mode (as appropriate) changes to any state other than QUIET within a time and voltage level specified in TBD.

Add the following after 83A.3.4.6 for the CAUI shutdown:

83A.3.4.7 Global energy detect function

The global energy detect function is mandatory for EEE capability. 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 may be considered a subset of the signal indication logic. If no energy is being received on the CAUI for the ingress direction then rx_mode shall be assigned the value QUIET; if no energy is being received on the CAUI for the egress direction then tx_mode shall be assigned the value QUIET. In all other situations, the values of rx_mode and tx_mode remain as communicated from the adjacent sublayers across the inter sublayer interfaces.

83A.4 Protocol implementation conformance statement (PICS) proforma for Annex 83A, 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)[13]

Editor's note (to be removed prior to final publication):

The PICS proforma will be updated when the content of this clause stabilizes.

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

Annex 83C

(informative)

PMA sublayer partitioning examples

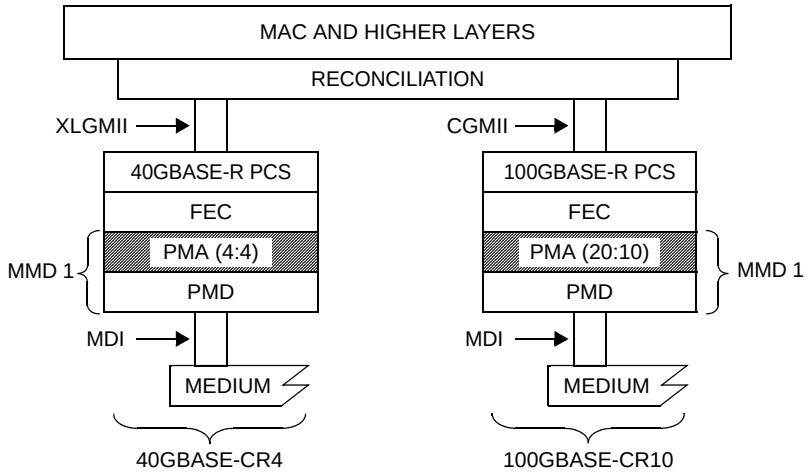
The following subclauses provide various partitioning examples. Partitioning guidelines and MMD numbering conventions are described in 83.1.4.

Change 83C.1 as follows:

83C.1 Partitioning examples with FEC

The example of BASE-R FEC (See Clause 74) implemented in a separate device from either the PCS or the PMD is illustrated in Figure 83–2.

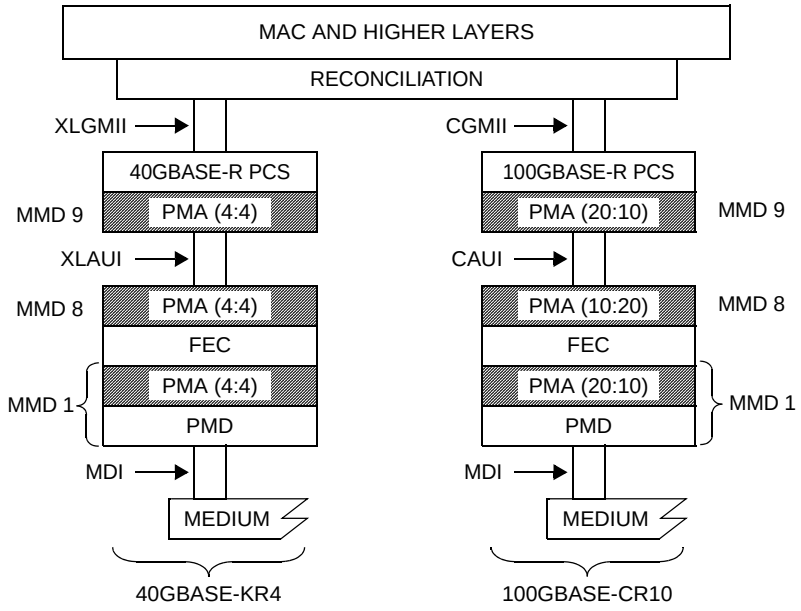
83C.1.1 FEC implemented with PCS



CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

Figure 83C–1—Example FEC implemented with PCS

83C.1.2 FEC implemented with PMD



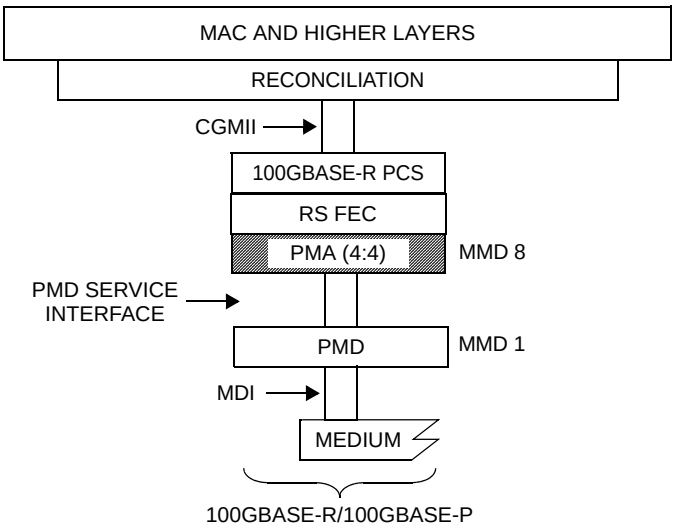
CAUI = 100 Gb/s ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLAUI = 40 Gb/s ATTACHMENT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

Figure 83C–2—Example FEC implemented with PMD

Add 83C.1a after 83C.1 as follows:

83C.1a Partitioning examples with RS FEC

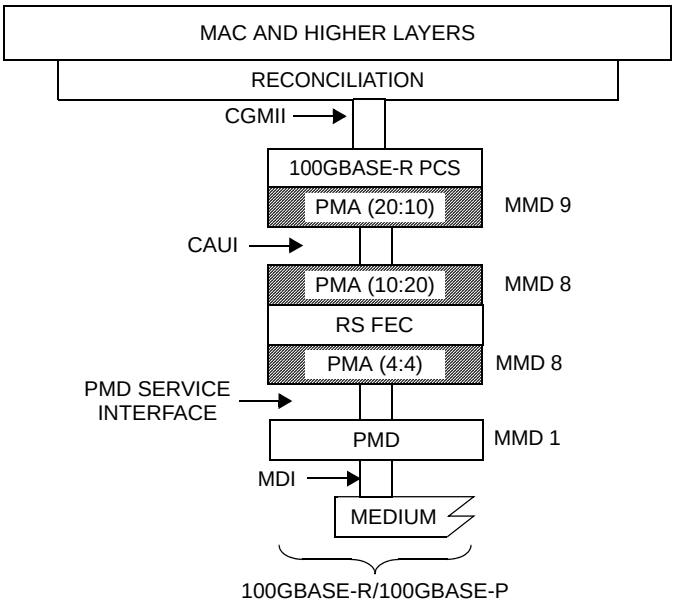
83C.1a.1 Single PMA sublayer with RS FEC



CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 83C–2a—Example single PMA sublayer with RS FEC

83C.1a.2 Single CAUI with RS FEC



CAUI = 100 Gb/s ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

Figure 83C–2a—Example single XLAUI/CAUI with RS FEC

Annex 92A

(informative)

100GBASE-CR4 TP0 and TP5 test point parameters

92A.1 Overview

Annex 92A provides information on parameters associated with test points TP0 and TP5 that may not be testable in an implemented system. TP0 and TP5 test points are illustrated in the 100GBASE-CR4 link block diagram of Figure 92–2.

92A.2 Transmitter characteristics at TP0

The transmitter characteristics at TP0 are defined in 93.8.1.

92A.3 Receiver characteristics at TP5

The receiver characteristics at TP5 are defined in 93.8.2.

92A.4 Transmitter and receiver differential printed circuit board trace loss

With the insertion loss TP0 to TP2 or TP3 to TP5 given in 92.8.3.4 and an assumed mated connector loss of 1.69 dB, the maximum insertion loss allocation for the transmitter and receiver differential controlled impedance printed circuit boards for each differential lane (i.e., the maximum value of the sum of the insertion losses from TP0 to the MDI host receptacle and from TP5 to the MDI host receptacle) are determined using Equation (92A–1). The maximum insertion loss allocation for the transmitter and receiver differential controlled impedance printed circuit boards is 13.62 dB at 12.9806 GHz. The maximum insertion loss for the transmitter or the receiver differential controlled impedance printed circuit board is one half of the maximum insertion loss $IL_{\mathrm{PCBmax}}(f)$.

Editor's note (to be removed prior to final publication):

The maximum insertion loss allocation for the transmitter or receiver PCB is $0.5 \times IL_{\mathrm{PCBmax}}$ = 6.81 dB at 12.8906 GHz.

$$IL_{\mathrm{PCB}}(f) \le IL_{\mathrm{PCBmax}}(f) \;=\; 0.0694 + 0.4248\sqrt{f} + 0.9322f \text{ (dB)} \qquad (92A\text{–}1)$$

for 0.01 GHz ≤ f ≤ 18.75 GHz.

where

f	is the frequency in GHz
$IL_{\mathrm{PCB}}(f)$	is the insertion loss for the transmitter and receiver PCB
$IL_{\mathrm{PCBmax}}(f)$	is the maximum insertion loss for the transmitter and receiver PCB

The minimum insertion loss allocation for the transmitter and receiver differential controlled impedance printed circuit boards for each differential lane (i.e., the minimum value of the sum of the insertion losses from TP0 to MDI receptacle and TP5 to MDI receptacle) are determined using Equation (92A–2). The minimum insertion loss for the transmitter or the receiver differential controlled impedance printed circuit board is one half of the minimum insertion loss $IL_{\mathrm{PCBmin}}(f)$.

$$IL_{\mathrm{PCB}}(f) \geq IL_{\mathrm{PCBmin}}(f) = 0.184(0.0694 + 0.4248\sqrt{f} + 0.9322f) \quad \text{(dB)} \qquad \text{(92A–2)}$$

for 0.01 GHz ≤ f ≤ 18.75 GHz.

where

f	is the frequency in GHz
$IL_{\mathrm{PCB}}(f)$	is the insertion loss for the transmitter and receiver PCB
$IL_{\mathrm{PCBmin}}(f)$	is the minimum insertion loss for the transmitter and receiver PCB

92A.5 Channel insertion loss

This subclause provides information on channel insertion losses for intended topologies ranging from 0.5 m to 5 m in length. The maximum channel insertion loss associated with the 5 m topology is determined using Equation (92A–3). The channel insertion loss associated with the 0.5 m topology and a maximum host channel is determined by Equation (92A–4). The channel insertion loss budget at 12.8906 GHz for the 5 m topology is illustrated in Figure 92A–1.

The maximum channel insertion loss for the 5 m topology is determined using Equation (92A–3). The maximum channel insertion loss is 35 dB at 12.8906 GHz.

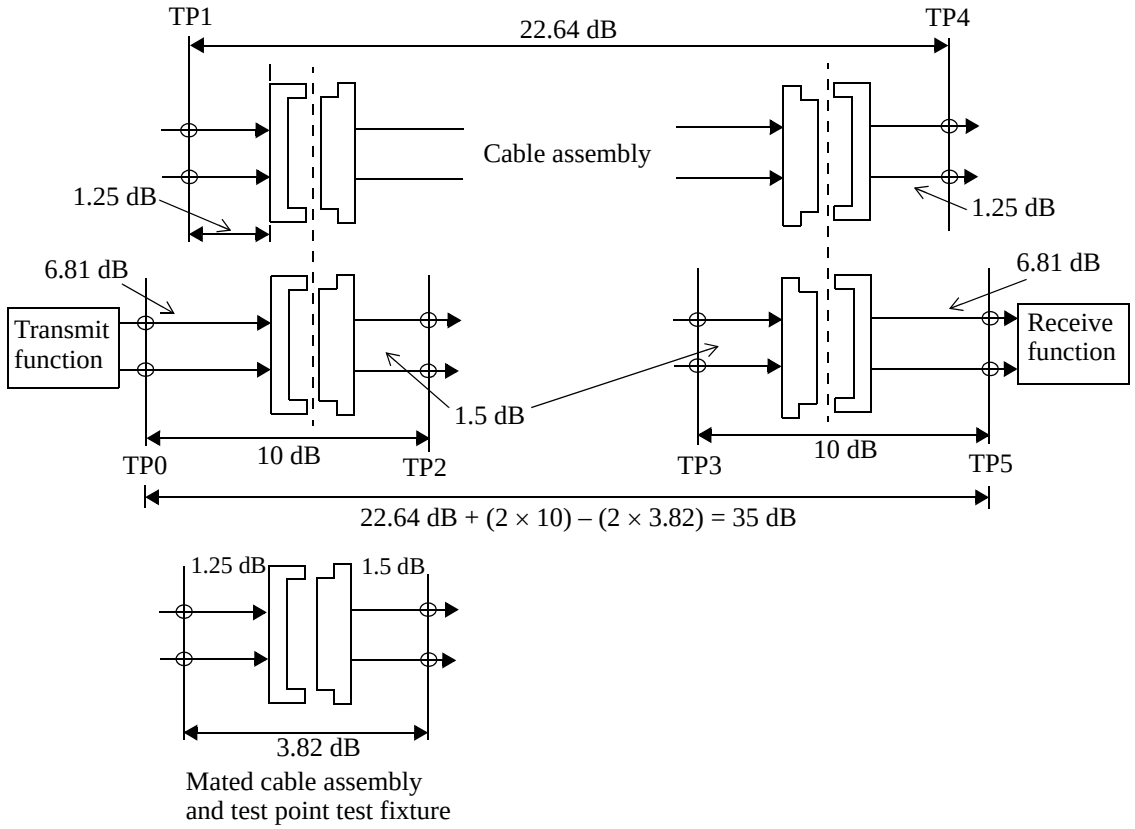
$$IL_{\mathrm{Chmax35dB}}(f) = IL_{\mathrm{Camax5m}}(f) + 2IL_{\mathrm{Host}}(f) - 2IL_{\mathrm{MatedTF}}(f) \text{ (dB)} \qquad \text{(92A–3)}$$

for 0.05 GHz ≤ f ≤ 18.75 GHz.

where

f	is the frequency in GHz
$IL_{\mathrm{Chmax35dB}}(f)$	is the maximum channel insertion loss between TP0 and TP5 representative of a 5 m cable assembly and a maximum host channel
$IL_{\mathrm{Camax5m}}(f)$	is the maximum 5 m cable assembly insertion loss using Equation (92–18)
$IL_{\mathrm{Host}}(f)$	is the maximum insertion loss from TP0 to TP2 or TP3 to TP5 using Equation (92–14)
$IL_{\mathrm{MatedTF}}(f)$	is the maximum insertion loss of the mated test fixture using Equation (92–35)

The channel insertion loss between TP0 and TP5 representative of a 0.5 m cable assembly and a maximum host channel is determined using Equation (92A–4).



NOTE—The connector insertion loss is 1.07 dB for the mated test fixture. The host connector is allocated 0.62 dB of additional margin.

Figure 92A–1—35 dB channel insertion loss budget at 12.8906 GHz

$$IL_{Ch0.5m}(f) = TBD \times IL_{Camax}(f) + 2IL_{Host}(f) - 2IL_{MatedTF}(f) \text{ (dB)} \quad (92A–4)$$

for 0.05 GHz ≤ f ≤ 18.75 GHz.

where

f	is the frequency in MHz
$IL_{Ch0.5m}(f)$	is the channel insertion loss between TP0 and TP5 representative of a 0.5 m cable assembly and a maximum host channel
$IL_{Camax}(f)$	is the maximum cable assembly insertion loss from TP1 to TP4 using Equation (92–18)
$IL_{Host}(f)$	is the maximum insertion loss from TP0 to TP2 or TP3 to TP5 using Equation (92–14)
$IL_{MatedTF}(f)$	is the maximum insertion loss of the mated test fixture using Equation (92–35)

92A.6 Channel return loss

The return loss of each lane of the 100GBASE-CR4 channel is recommended to meet the values determined using Equation (92–24).

92A.7 Channel insertion loss deviation (ILD)

The channel insertion loss deviation is the difference between the channel insertion loss and the fitted channel insertion loss is determined using Equation (92A–5).

$$ILD(f) = IL_{\mathrm{Ch}}(f) - IL_{\mathrm{fitted}}(f) \text{ (dB)} \qquad \text{(92A–5)}$$

where

f	is the frequency in GHz
$ILD(f)$	is the channel insertion loss deviation at frequency f
$IL_{\mathrm{fitted}}(f)$	is defined in Equation (92–18)

Given the channel insertion loss is at N uniformly-spaced frequencies f_n spanning the frequency range 0.05 GHz to 18.75 GHz with a maximum frequency spacing of 0.01 GHz, the coefficients of the fitted channel insertion loss are determined using Equation (92–19) and Equation (92–20).

The channel insertion loss deviation is recommended to be within the region defined by Equation (92–22) and Equation (92–23) for all frequencies from 0.05 GHz to 18.75 Gw3qw0-Hz.

92A.8 Channel integrated crosstalk noise (ICN)

Since four lanes are used to transfer data between PMDs, the near-end crosstalk (NEXT) that is coupled into a victim receiver will be from the four adjacent transmitters. The channel multiple disturber NEXT loss, *MDNEXT_loss*(*f*), is specified using the individual NEXT losses as shown in Equation (92–25).

In addition, the far-end Crosstalk (FEXT) that is coupled into a receiver will be from the three other transmitters adjacent to the victim transmitter. The channel multiple disturber FEXT loss, *MDFEXT_loss*(*f*), is specified using the individual FEXT losses as shown in Equation (92–26).

Given the channel *MDNEXT_loss*(*f*) and *MDFEXT_loss*(*f*) measured over N uniformly-spaced frequencies f_n spanning the frequency range 50 MHz to 20000 MHz with a maximum frequency spacing of 10 MHz, the RMS value of the integrated crosstalk noise is determined using Equation (92–27) through Equation (92–31) and the parameters shown in Table 92–11.

The total integrated crosstalk RMS noise voltage of the channel is recommended to meet the values determined using Equation (92A–6) illustrated in Figure 92A–2.

$$\sigma_{x,\,ch} \le \{ TBD \} \qquad (\mathrm{mV}) \qquad \text{(92A–6)}$$

where *IL* is the value of the channel insertion loss in dB at 12.8906 GHz.

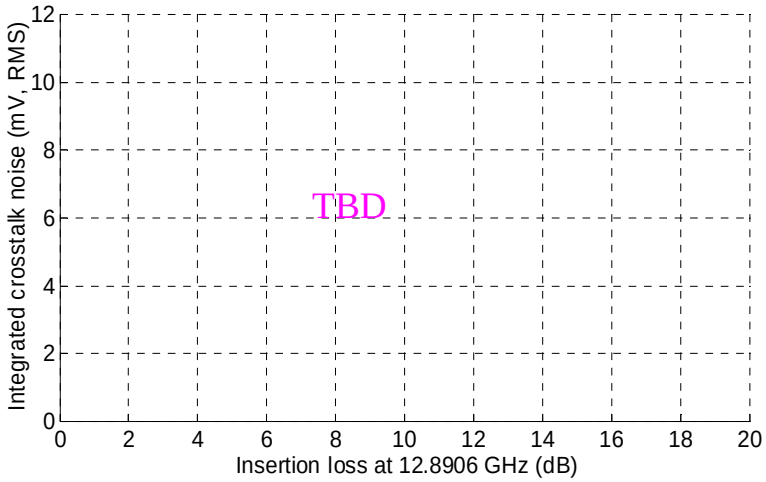


Figure 92A–2—Channel integrated crosstalk noise

Annex 93A

(normative)

Characteristics of electrical backplanes

93A.1 Channel operating margin

Editor's note (to be removed prior to final publication):

This annex contains a mathematical description of the channel operating margin computation. It is expected that a sample implementation of this calculation, in MATLAB(R) or similar computational language, will be included as a separate informative annex when the content of this annex stabilizes.

The channel operating margin (COM) is a figure of merit for a channel derived from a measurement of its scattering parameters. The channel operating margin is related to the ratio of a calculated signal amplitude to a calculated noise amplitude as defined by Equation (93A–1).

$$COM = 20\log_{10}(A_s/A_n) - COM_0 \quad \text{(93A–1)}$$

The signal amplitude A_s is defined in 93A.1.5 and the noise amplitude A_n is defined in 93A.1.6. A minimum channel operating margin is allocated using the term COM_0. This term accounts for transmitter and receiver implementation considerations and channel impairments that are not explicitly included in the analysis.

The parameters used to calculate COM are summarized in Table 93A–1. The values assigned to these parameters are defined by the Physical Layer specification that invokes the method.

93A.1.1 Measurement of the channel

The channel consists of a victim signal path plus some number of far-end and near-end crosstalk paths. The total number of paths for a given channel is denoted as K and, by convention, the path index k=0 corresponds to the victim path. The number of crosstalk paths is a function of the structure of the system. All significant contributors to the channel crosstalk should be included in the calculation of COM.

Each signal path is represented by a set of frequency-dependent scattering parameters. For the purpose of the calculation of COM, references to scattering parameters correspond to the differential-mode scattering parameters. The scattering parameters measured at frequency f are presented as the 2 x 2 matrix $S(f)$ as defined by Equation (93A–2).

$$S(f) = \begin{bmatrix} s_{11}(f) & s_{12}(f) \\ s_{21}(f) & s_{22}(f) \end{bmatrix} \quad \text{(93A–2)}$$

The relationship between $S(f)$ and other commonly cited characteristics is as follows. The insertion loss is the magnitude in dB of either $1/s_{12}(f)$ or $1/s_{21}(f)$. The input and output return loss is the magnitude in dB of $1/s_{11}(f)$ and $1/s_{22}(f)$ respectively.

Table 93A–1—Summary of parameters

Parameter	Symbol	Units
Signaling rate	f_b	GHz
Maximum start frequency	$f_{\min}$	GHz
Maximum frequency step	Δf	GHz
Transmitter reflection coefficient DC value Reference frequency	Γ_1 Γ_{01} f_1	 V/V GHz
Receiver reflection coefficient DC value Reference frequency	Γ_2 Γ_{02} f_2	 V/V GHz
Transmitter differential peak output voltage Victim Far-end aggressor Near-end aggressor	 A_v A_f A_n	 V V V
Transmitter 3 dB bandwidth Victim Far-end aggressor Near-end aggressor	 f_v f_f f_n	 GHz GHz GHz
Receiver 3 dB bandwidth	f_r	
Transmitter equalizer, pre-cursor coefficient Minimum value Maximum value Step size	$c(-1)$	 — — —
Transmitter equalizer, post-cursor coefficient Minimum value Maximum value Step size	$c(1)$	 — — —
Continuous time filter, DC gain Minimum value Maximum value Step size	g_{DC}	 dB dB dB
Number of signal levels	L	—
Number of samples per unit interval	M	—
Victim single bit response exception window	W	UI
Normalized RMS Gaussian noise	σ_G	—
Normalized peak dual-Dirac noise	A_{DD}	—
Target uncorrected symbol error ratio	SER_0	—
Minimum channel operating margin	COM_0	dB

The scattering parameters for the victim signal path are measured from TP0 to TP5. The scattering parameters for each crosstalk path are measured from the package-to-board interface of the aggressor transmitter to TP5. The frequency-dependent scattering matrix for signal path k is denoted as $S^{(k)}(f)$. The reference impedance for scattering parameter measurements is 100 Ω.

It is recommended that the scattering parameters be measured with uniform time step no larger than Δf from a start frequency no larger than $f_{\min}$ to a stop frequency of at least the signaling rate f_b.

93A.1.2 Path terminations

The input to each signal path is terminated by an impedance defined by the reflection coefficient Γ_1. The output of each signal path is terminated by an impedance defined by the reflection coefficient Γ_2.

Editor's note (to be removed prior to final publication):

The transmitter and receiver termination models were stated to be TBD in mellitz_01_0712.pdf. The following equations were used as placeholders in that presentation.

The reflection coefficient Γ_1 is defined by Equation (93A–3) and the reflection coefficient Γ_2 is defined by Equation (93A–4) where $j = \sqrt{-1}$.

$$\Gamma_1 = \frac{\Gamma_{01} + j(f/f_1)}{1 + j(f/f_1)} \quad \text{(93A–3)}$$

$$\Gamma_2 = \frac{\Gamma_{02} + j(f/f_2)}{1 + j(f/f_2)} \quad \text{(93A–4)}$$

The voltage transfer function of the terminated signal path is defined by Equation (93A–5) where $\Delta S(f) = s_{11}(f)s_{22}(f) - s_{12}(f)s_{21}(f)$.

$$H_{21}(f) = \frac{s_{21}(f)}{1 - s_{11}(f)\Gamma_1(f) - s_{22}(f)\Gamma_2(f) + \Gamma_1(f)\Gamma_2(f)\Delta S(f)} \quad \text{(93A–5)}$$

The voltage transfer function for the signal path represented by $S^{(k)}(f)$ is denoted $H_{21}^{(k)}(f)$.

93A.1.3 Filters

The voltage transfer function for each signal path $H_{21}^{(k)}(f)$ (see 93A.1.2) is multiplied by $H_t(f)$ and $H_r(f)$ to yield $H_{tr}^{(k)}(f)$.

$H_t(f)$ is defined by Equation (93A–6) and is intended to represent the rise and fall times of the transmitter. It is a function of the parameter f_t which is set based on the path index *k*. If *k*=0, i.e. the victim path, then $f_t = f_v$. If *k* corresponds to a far-end crosstalk path then $f_t = f_f$. If *k* corresponds to a near-end crosstalk path then $f_t = f_n$.

$$H_t(f) = \frac{A_t}{1 - (f/f_t)^2 + j1.414214(f/f_t)} \quad \text{(93A–6)}$$

$H_r(f)$ is a noise filter defined by Equation (93A–7).

$$H_r(f) = \frac{1}{1 - 3.414214(f/f_r)^2 + (f/f_r)^4 + j2.613126(f/f_r - (f/f_r)^3)} \quad \text{(93A–7)}$$

The resulting $H_{tr}^{(k)}(f)$ is then multiplied by $H_{ffe}(f)$ and $H_{ctf}(f)$ to yield $H^{(k)}(f)$.

$H_{ffe}(f)$ is defined by Equation (93A–8) and is intended to represent the transmitter equalizer. If k corresponds to a near-end crosstalk path, then $c(-1)$ and $c(1)$ are zero regardless of the values used for the other paths. The value of the "cursor" coefficient c(0) is set to $1 - |c(-1)| - |c(1)|$ for any value of $c(-1)$ and $c(1)$.

$$H_{ffe}(f) = \sum_{i=-1}^{1} c(i)\exp(-j2\pi(i+1)(f/f_b)) \quad \text{(93A–8)}$$

$H_{ctf}(f)$ is defined by Equation (93A–9).

$$H_{ctf}(f) = \frac{10^{g_{DC}/20} + j(4f/f_b)}{(1 + j(f/f_b))(1 + j(4f/f_b))} \quad \text{(93A–9)}$$

The filtered voltage transfer function $H^{(k)}(f)$ is used to compute the single bit response (see 93A.1.4).

93A.1.4 Single bit response

The single bit response of a signal path is defined to be the output of the path following the application of rectangular pulse one unit interval in duration at its input. First define the function $X(f)$ per Equation (93A–10) where $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ and $T_b = 1/f_b$ is the unit interval.

$$X(f) = A_t T_b \text{sinc}(fT_b)/(L-1) \quad \text{(93A–10)}$$

$X(f)$ is a function of A_t which in turn is based on the path index k. If k=0, i.e. the victim path, then $A_t = A_v$. If k corresponds to a far-end crosstalk path then $A_t = A_f$. If k corresponds to a near-end crosstalk path then $A_t = A_n$

The single bit response $h^{(k)}(t)$ is derived from the voltage transfer function $H^{(k)}(f)$ (see 93A.1.3) using Equation (93A–11).

$$h^{(k)}(t) = \int_{-\infty}^{\infty} X(f)H^{(k)}(f)\exp(j2\pi ft)dt \quad \text{(93A–11)}$$

NOTE 1—It is expected that COM will be computed from measurements at discrete frequencies that cover a limited span (see 93A.1.1). The inverse Fourier transform depicted in Equation (93A–11) will likely be implemented as a discrete Fourier transform and the filtered voltage transfer function may need to be extrapolated (both to DC and the Nyquist frequency) for this computation. The extrapolation method and Nyquist frequency must be chosen carefully to limit the error in the COM computation.

NOTE 2—The time span of the single bit response in unit intervals, N, is limited in practice by frequency step Δf ($N = f_b/\Delta f$) but in general should be set to include all significant components of the single bit response.

93A.1.5 Signal amplitude

Editor's note (to be removed prior to final publication):

This procedure was included in mellitz_01_0712.pdf as an example for a 100GBASE-KR4 PHY (PAM2 modulation). The procedure itself is TBD for 100GBASE-KR4 and 100GBASE-KP4.

COM is a function of the variables $c(-1)$, $c(1)$, and g_{DC}. The following procedure is used to determine the values of these variables that will be used to calculate COM.

a) Compute the single bit response of the victim path $h^{(0)}(t)$ for a given $c(-1)$, $c(1)$, and g_{DC} using the procedure defined in 93A.1.4.
b) Define t_z to be the zero crossing of the rising edge of the single bit response.
c) Define t_s to be t_z+T_b. Define A_s to be $h^{(0)}(t_s)$.
d) Force the single bit response to zero within the exception window defined as $[t_z, t_z+WT_b]$. The result is $h_w(t)$ as defined by Equation (93A–12). Compute the variance of $h_w(t)$ per Equation (93A–13).
e) Compute the figure of merit (FOM) per Equation (93A–14). Note that the additive Gaussian noise term σ_G has been included.

$$h_w(t) = \begin{cases} 0, & t_z \le t < t_z + WT_b \\ h^{(0)}(t), & \text{otherwise} \end{cases} \quad \text{(93A–12)}$$

$$\sigma_w^2 = \frac{1}{NT_b}\int_0^{NT_b} h_w^2(t)dt \quad \text{(93A–13)}$$

$$FOM = 10\log_{10}\left(\frac{A_s^2}{\sigma_w^2 + A_s^2\sigma_G^2}\right) \quad \text{(93A–14)}$$

The FOM is calculated for each permitted combination of $c(-1)$, $c(1)$, and g_{DC} values per Table 93A–1. The combination of values that maximizes the FOM, including the corresponding values of t_z and A_s, is used for the calculation of the interference and noise amplitude in 93A.1.6 and the calculation of COM in 93A.1.

93A.1.6 Interference and noise amplitude

Given the values of $c(-1)$, $c(1)$, g_{DC}, t_z, and A_s derived in 93A.1.5, compute the combined interference and noise distribution $p(y)$ per 93A.1.6.3. The corresponding cumulative distribution function is $P(y)$ as defined by Equation (93A–15).

$$P(y) = \int_{-\infty}^{y} p(y)dy \quad \text{(93A–15)}$$

The noise amplitude, A_n, is the magnitude of the value of y_0 that satisfies the relationship $P(y_0) = SER_0$ where SER_0 is the target uncorrected symbol error rate.

93A.1.6.1 Interference amplitude distribution

The interference amplitude distribution is computed from the single bit response $h(t)$ with the assumption that the transmitted symbols are independent, identically distributed random variables and that the symbols are uniformly distributed across the set of L possible values. For the purpose of this subclause, $h(t)$ is a general notation that corresponds to either the windowed victim single bit response $h_w(t)$ or the single bit response $h^{(k)}(t)$ corresponding to the kth crosstalk path (see 93A.1.6.3).

Sample the single bit response with uniform time step T_b starting at time $t_m = mT_b/M$ as shown in Equation (93A–16).

$$h_m(n) = h(t_m + nT_b) \quad \text{(93A–16)}$$

The phase index *m* can assume any integer value from 0 to *M*–1. The symbol index *n* can assume any integer value from 0 to *N*–1 where *N* is the time span of the single bit response in unit intervals.

Equation (93A–17) defines the *n*th component of the interference amplitude distribution function where $\delta(y)$ is the Dirac delta function.

$$p_{m,n}(y) = \frac{1}{L}\sum_{l=0}^{L-1}\delta\left(y-\left(\frac{2l}{L-1}-1\right)h_m(n)\right) \quad (93A–17)$$

The set of *N* such components are combined via convolution to obtain the complete interference amplitude distribution. For a given phase index *m*, initialize $p_m(y)$ to 1 for y=0 (0 otherwise) and then evaluate Equation (93A–18) sequentially for n=0 to *N*–1.

$$p_m(y) = p_m(y) * p_{m,n}(y) \quad (93A–18)$$

In Equation (93A–18), "*" denotes convolution which is defined by Equation (93A–19).

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \quad (93A–19)$$

NOTE—It is expected that COM will be numerically computed using a quantized amplitude axis *y*. The amplitude step Δy will introduce quantization error in the calculated distribution function that will be compounded by subsequent convolutions with other quantized distribution functions. It is recommended that Δy be no larger than 0.1 mV in order to limit the error.

The variance of the interference amplitude for phase index *m* is computed from the amplitude distribution using Equation (93A–20).

$$\sigma_m^2 = \int_{-\infty}^{\infty} y^2 p_m(y)dy \quad (93A–20)$$

The interference amplitude distribution is computed for each phase index m=0 to *M*–1. The value of *m* that yields the largest variance is used for the calculation of COM. The interference amplitude distribution that corresponds to this value of *m* is denoted as $p(y)$.

93A.1.6.2 Noise amplitude distribution

The calculation of COM includes two additive noise terms that are described in terms of their distribution function. The first additive noise component has a Gaussian noise amplitude distribution function with zero mean and standard deviation $A_s\sigma_G$ where A_s is the signal amplitude computed per 93A.1.5. The amplitude distribution of the additive Gaussian noise term is defined by Equation (93A–21).

$$p_G(y) = \frac{1}{A_s\sigma_G\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{y}{A_s\sigma_G}\right)^2\right) \quad (93A–21)$$

The second component of additive noise has a dual-Dirac distribution with zero mean and peak amplitude A_sA_{DD}. The amplitude distribution of the additive dual-Dirac noise is defined by Equation (93A–22).

$$p_{DD}(y) = \frac{1}{2}\delta(y-A_sA_{DD}) + \frac{1}{2}\delta(y+A_sA_{DD}) \quad (93A–22)$$

The components are combined using convolution to yield the overall noise amplitude distribution function as defined in Equation (93A–23).

$$p_n(y) = p_G(y) * p_{DD}(y) \quad (93A–23)$$

93A.1.6.3 Combination of interference and noise distributions

Editor's note (to be removed prior to final publication):

The victim path processing described below was included in mellitz_01_0712.pdf as an example for a 100GBASE-KR4 PHY (PAM2 modulation). The procedure is TBD for 100GBASE-KR4 and 100GBASE-KP4.

For the victim path, represented by the single bit response $h^{(0)}(t)$, components outside of the exception windows defined by $[t_z, t_z+WT_b)$ are considered interference. Define $h_w(t)$ using Equation (93A–12) and compute the distribution of interference amplitudes using the procedure defined in 93A.1.6.1. Denote the result as $p(y)$.

The contributions of the K–1 crosstalk paths to the total interference are included as follows. Given the single bit response of kth crosstalk path, $h^{(k)}(t)$, compute the interference amplitude distribution using the procedure defined in 93A.1.6.1, denote the result as $p^{(k)}(y)$, and evaluate Equation (93A–24). Repeat sequentially for integer values k=1 to K–1.

$$p(y) = p(y) * p^{(k)}(y) \quad (93A–24)$$

The additive noise distribution defined in 93A.1.6.2, $p_n(y)$, is then included to yield the combined interference and noise amplitude distribution as shown in Equation (93A–25).

$$p(y) = p(y) * p_n(y) \quad (93A–25)$$

Annex 93B

(informative)

Electrical backplane reference model

This annex describes additional informative test points that may used to partition the electrical backplane channel.

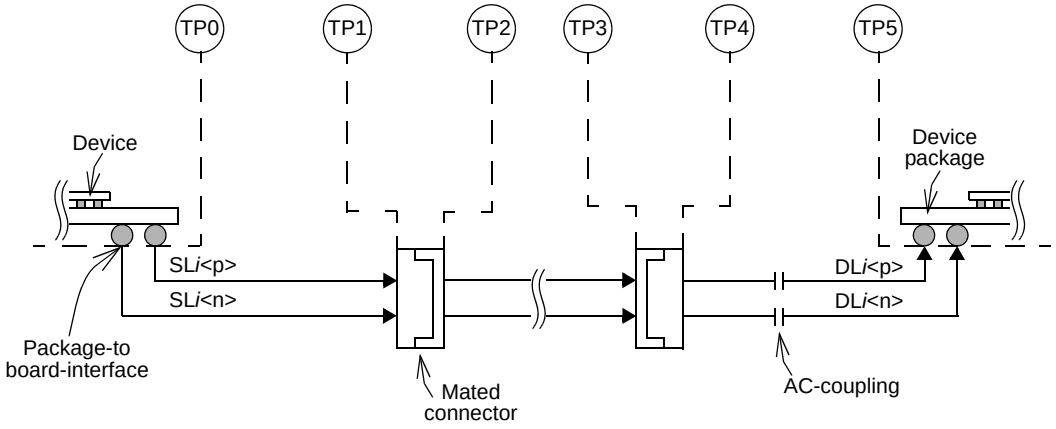


Figure 93B–1—Reference model (one direction from one lane is illustrated)

Table 93B–1—Description of channel components

Test points	Description
TP1 to TP1	The printed circuit board between the transmitter and the separable connector closest to the transmitter. TP1 is defined to be the interface between the board and connector plug.
TP2 to TP3	The electrical path from the separable connector closest to the transmitter to the separable connector closest to the receiver. TP2 and TP3 are defined to be the interface between connector receptacle and the printed circuit board.
TP4 to TP5	The printed circuit board between the receiver and the separable connector closest to the receiver. TP4 is defined to be the interface between the board and connector plug. It is recommended that the AC-coupling capacitors are implemented between TP4 and TP5
TP0 to TP5	The electrical backplane channel as defined in 93.9 and 94.4. TP0 and TP5 are defined to be the interface between the device package and the printed circuit board.